

Vijay Singh Rathore ·  
Subhash Chander Sharma ·  
Joao Manuel R. S. Tavares ·  
Catarina Moreira ·  
B. Surendiran *Editors*

# Rising Threats in Expert Applications and Solutions

Proceedings of FICR-TEAS 2022

# **Lecture Notes in Networks and Systems**

**Volume 434**

## **Series Editor**

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

## **Advisory Editors**

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

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

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

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

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

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

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

Imre J. Rudas, Óbuda University, Budapest, Hungary

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

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

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

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

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

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

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

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

More information about this series at <https://link.springer.com/bookseries/15179>

Vijay Singh Rathore · Subhash Chander Sharma ·  
Joao Manuel R. S. Tavares · Catarina Moreira ·  
B. Surendiran  
Editors

# Rising Threats in Expert Applications and Solutions

Proceedings of FICR-TEAS 2022



Springer

*Editors*

Vijay Singh Rathore  
IIS (Deemed to be University)  
Jaipur, Rajasthan, India

Subhash Chander Sharma  
Electronics and Computer Discipline  
Indian Institute of Technology  
Roorkee, Uttarakhand, India

Joao Manuel R. S. Tavares   
Faculdade de Engenharia  
Universidade do Porto  
Porto, Portugal

Catarina Moreira  
Faculty of Science  
School Information Systems  
Queensland University of Technology  
Brisbane, QLD, Australia

B. Surendiran  
Department of Computer Science  
and Engineering  
National Institute of Technology  
Puducherry  
Karaikal, Pondicherry, India

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-19-1121-7

ISBN 978-981-19-1122-4 (eBook)

<https://doi.org/10.1007/978-981-19-1122-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

# Preface

The 2nd FICR International Conference on Rising Threats in Expert Applications and Solutions (FICR-TEAS 2022) was held at Jaipur, India, during 7–8 January 2022 in hybrid mode. The conference was organized by the “Department of Computer Science and Information Technology” and “International Cell IIS-TIE” at IIS (deemed to be University), Jaipur, and supported by Science and Engineering Research Board (SERB), DST, Government of India. The 1st FICR-TEAS International Conference “FICR-TEAS 2020” was successfully organized on 17–19 January 2020 at IIS (deemed to be University), Jaipur, in association with Springer Nature for publication (AISC Series) and supported by CSI, ACM, IEEE Ireland & UK Section. The conference addressed recent technological developments, specifically the “expert applications” followed by the analysis of the Rising Threats associated with them and eventually proposing solutions to these threats.

COVID-19 knows no boundaries or borders—the transfer of knowledge should not either. Technology has transformed with great speed in the last few decades, resulting in the development of expert applications making life more effortless. The conference raised awareness about issues related to increase in threats in expert applications and solutions, which will aid in the creation of better solutions for the society. The COVID-19 pandemic has impacted us more than any other event in most of our lifetimes. Companies, associations, and destinations globally are trying to navigate their way through this crisis, balancing the short-term need with a long-term strategy. While we are all in the same storm, we must realize that we are in different boats, and therefore, different solutions and strategies are necessary.

To understand another dimension of the conference abbreviation, it would be better to understand the word “FICR”, Indian word (pronounced as “Ficr”), which means “Seeking Solutions For Worries”, i.e. giving detailed thoughts to a particular issue by analysing all its pros and cons in advance, so that the issues can be addressed with proper planning and utmost care to benefit the concerned. Also, TEAS represents the most popular hot beverage “Tea” all over the world, which brings freshness. Here, through the conference “FICR-TEAS 2022”, the “worries” (Ficr) of the rising threats in the expert applications would be discussed, analysed, and probably solved, during various tea sessions (and tea breaks) of the conference.

FICR-TEAS 2022 was organized keeping these dimensions at preference. The conference aimed to provide an international platform to the researchers, academicians, industry representatives, government officials, students, and other stakeholders in the field to explore the opportunities, to disseminate and acquire beneficial knowledge from the various issues deliberated in the paper presented on different themes in the conference. The technical program committee and advisory board of FICR-TEAS 2022 included eminent academicians, researchers, and practitioners from abroad as well as from all over the nation.

The conference received incredible response from both delegates and students in reference to research paper presentations. More than 256 papers were received, out of which 75 were selected after impartial plagiarism check and rigorous peer-review process. In all, 80 oral presentations were delivered, and six posters were presented. Paper presenters came to attend the conference from all corners of the country. We had international participants and delegates as well from countries like Italy, Nepal, Portugal, Romania, USA, Australia, Belgrade (Serbia), and Poland, to name a few.

We are deeply appreciative towards all our authors for having shown confidence in us and considering FICR-TEAS 2022 a platform for sharing and presenting their original research work. We also express our sincere gratitude to the focused team of chairs, co-chairs, international advisory committee, and technical program committee. We are also gratified to Mr. Aninda Bose, (Senior Publishing Editor, Springer Nature, India) for providing continuous guidance and support. Also, we extend our heartfelt thankfulness and appreciation towards the reviewers and technical program committee members for showing their concern and efforts in the review process. We are indeed thankful to everyone associated directly or indirectly with the conference, organizing a firm team and leading it towards success.

We hope you enjoy the conference proceedings and wish you all the best.

Jaipur, India

Prof. Dr. Vijay Singh Rathore  
Prof. Joao Manuel R. S. Tavares  
Prof. Subhash Chander Sharma  
Dr. Catarina Moreira  
Dr. B. Surendiran

# Acknowledgements

The organization of a successful conference always depends upon the support received during its planning and execution.

At the outset, we are grateful to Dr. Ashok Gupta, Hon'ble Chancellor IIS (deemed to be university) for his unstinted support and invaluable guidance throughout, which gave this three-day event its present shape. We are also grateful to Prof. T. N. Mathur, Vice Chancellor IIS (deemed to be university), and Prof. Raakhi Gupta, Rector and Registrar, IIS (deemed to be university) for the meticulous planning and attention to details which has helped us in organizing this event. Our sincere thanks are also due to Prof. K. S. Sharma, Advisor, and Prof. Pradeep Bhatnagar, Dean IIS (deemed to be university), for their kind support and advice since the conception of this idea.

We are grateful to a number of people without whom we would not have been able to successfully organize this mega event and are personally thankful to Prof. R. S. Salaria, Guru Nanak Technical Institutions, Hyderabad; Dr. Yogesh Singh Chauhan, IIT, Kanpur; Prof. Subhash Chander Sharma IIT, Roorkee, India; Prof. Mike Hinchey, Chair, University of Limerick, Ireland; Mr. Aninda Bose, Springer Nature, Germany; Prof. Joao Manuel R. S. Tavares, University of Porto, Portugal; Prof. Milan Tuba, Singidunum University, Belgrade, Serbia; Prof. Rosalina Babo, Polytechnic Institute of Porto, Portugal; Prof. Vincenzo Piuri, University of Milan, Italy; Prof. Zdzislaw Polkowski, Jan Wyzykowski University, Poland; Mr. Peter Kent, UKEI, London, UK; Dr. Catarina Moreira, Queensland University of Technology, Brisbane, Australia; Dr. B. Surendiran, NIT Karaikal, Pondicherry, India; Prof. M. N. Hoda, Bharti Vidyapeeth University, New Delhi; Prof. Valentina Emilia Balas, Aurel Vlaicu, University Of Arad, Romania; Dr. Dharm Singh, Namibia University of Science and Technology, Namibia; Prof. (Dr.) M. Hanumanthappa, Bangalore University, Bengaluru; Dr. Satish Kumar Singh, IIIT Allahabad; Prof. (Dr.) Durgesh Kumar Mishra, Sri Aurobindo Institute of Technology, Indore; Prof. Vipin Tyagi, Computer Society of India; Dr. Nilanjan Dey, JIS University, Kolkata; Dr. Niko Phillips, International, Activate Learning, City of Oxford College, Oxford, UK; Prof. Cornelia-Victoria Anghel Drugărin, Babeş-Bolyai University Cluj-Napoca, România; Dr. Dana Rad, Aurel Vlaicu, University of Arad, Arad, Romania; Dr. Daniele Ritelli, University of Bologna, Bologna, Italy; Dr. Marina Popolizio,

Department of Electrical And Information Engineering, Polytechnic of Bari, Italy; Dr. Vincenzo Di Lecce, Department of Electrical And Information Engineering, Polytechnic of Bari, Italy; Dr. Nhu Gia Nguyen, Duy Tan University, Da Nang, Vietnam; Prof. V. S. Sundararajan, Ministry of Singapore, Singapore; Dr. Kishor Sarawadekar, IIT BHU, Varanasi; Prof. P. K. Mishra, Institute of Science, BHU, Varanasi; Prof. Maya Ingle, DAVV, Indore; Prof. N. N. Jani, Saurashtra University, Rajkot; Prof. Reena Dadhich, Department of Computer Science, University of Kota, Kota; Prof. (Dr.) M. Sundaresan, Department of IT, Bharthiar University, Coimbatore; Dr. O. P. Rishi, University of Kota, Kota; Dr. M. P. Singh, B. R. Ambedkar University, Agra; Prof. Vibhakar Mansotra, University of Jammu, Jammu; Dr. P. V. Virparia, Sardar Patel University, Gujarat; Dr. Neeraj Bhargava, MDS University, Ajmer; Prof. Pawanesh Abrol, University of Jammu, Jammu; Dr. S. Kartihik, SNS Engineering College, Coimbatore; Prof. C. K. Kumbharana, Saurashtra University, Rajkot; Dr. B. L. Shivakumar, Sri Ramakrishna Polytechnic College, Coimbatore; Dr. Atul Gonsai, Saurashtra University, Rajkot; Dr. Dhiren Patel, Gujarat Vidyapeeth University, Ahmedabad; Prof. Naseeb S. Gill, MDU, Rohtak; Dr. N. K. Joshi, MIMT, Kota; Prof. Vinod Sharma, University of Jammu, Jammu; Dr. Tanupriya Chaudhary, UPES, Dehradun; Dr. Praveen Kumar Vashishth, Amity University, Noida; Dr. Bankim Patel, U. T. University, Baroda; Dr. Meenakshi Tripathi, MNIT, Jaipur; Dr. Amit Joshi, GR Foundation, Ahmedabad; Dr. Jatinder Manhas, University of Jammu, Jammu; Prof. Madhavi Sinha, BITS, Jaipur; Dr. Mayank Agrawal, Gurukul Kangri University, Haridwar; Dr. Ashutosh Saxena, University of Hyderabad Campus, Hyderabad; Dr. Ritu Bhargava, Sophia College, Ajmer; Dr. Kunwar Singh Vaisla, BTKIT, Dwarahat, Uttrakhand; Dr. Deepshikha Bhargava, UPES, Dehradun; Dr. Nishtha Keswani, CUoR, Kishangarh; Dr. Vivek Tiwari, IIIT Naya Raipur, Chhattisgarh; Dr. Sumeet Gill, MDU, Rohtak; Mr. Mihir Chauhan, GR Foundation, Ahmedabad; Dr. Suresh Kumar, Savitha Engineering College, Chennai; Dr. Avinash Panwar, MLSU, Udaipur; Dr. Avinash Sharma, MMEC, Haryana; Dr. Pritee Parwekar, SRM Institute of Science and Technology, Ghaziabad; Dr. Pinal J. Patel, GEC, Gandhinagar; Dr. Lavika Goel, MNIT, Jaipur; Dr. N Suganthi, SRM University, Chennai; Dr. Vikrant Bhateja, SRMGP, Lucknow; Dr. Mahipal Singh Deora, B. N. University, Udaipur; Dr. Abid Hussain, Career Point University, Kota; Dr. Sonali Vyas, UPES, Dehradun; Dr. Seema Rawat, Amity University, Tashkent, Uzbekistan; Dr. Minakshi Memoria, UIT, Uttaranchal; Dr. Amit Sharma, Career Point University, Kota; Dr. Bharat Singh Deora, JRN Rajasthan Vidyapeeth University, Udaipur; Ms. Gunjan Gupta, A2G Enterprises Pvt. Ltd., Noida; Prof. Krishna Gupta, University of Rajasthan, Jaipur; Dr. Paras Kothari, CSI, Udaipur Chapter; Prof. Ashok Agrawal, University of Rajasthan, Jaipur; Dr. Salini Suresh, Dayananda Sagar Institutions, Bangalore; Dr. Shikha Maheshwari, Manipal University, Jaipur; Prof. Swati V. Chande, ISIM, Jaipur; Mr. Vipin Jain, Metacube Software Pvt. Ltd., Jaipur; Dr. Manju Mandot, Maharishi Dayanand University, Rohtak, and all other resource persons, experts, and guest speakers who so graciously consented to deliver the invited talks.

We are deeply appreciative towards all our authors for having shown confidence in us and considering FICR-TEAS 2022 a platform for sharing and presenting their

original research work. We also express our sincere gratitude to the focused team of chairs, co-chairs, reviewers, international advisory committee, and technical program committee.

We are also thankful to Shri. Aninda Bose—Senior Editor, Hard Sciences, Springer, for providing continuous guidance and support.

No task in this world can be completed successfully without the support of your team members. We would like to extend our heartfelt appreciation to Dr. Anubha Jain, Head (CS and IT), IIS (deemed to be university), and Organizing Secretary, FICR-TEAS 2022; Dr. Neha Tiwari and Dr. Navneet Sharma, Co-convenor, FICR-TEAS 2022, for their keen support and novel suggestions. Thanks is also due to Dr. Ruchi Nanda and Dr. Amita Sharma, Co-Organizing Secretary, FICR-TEAS 2022 for their contribution.

We would also like to record our gratitude to our team members Dr. Deepshikha Bhatia, Dr. Priyanka Verma, Mr. Rajneesh Chaturvedi, Dr. Astha Pareek, Ms. Harshita Bhargava, Ms. Priyanka Dewatwal, Mr. Tikesh Kumawat, Mr. Bhaskar Mani Tiwari, Ms. Kriti Manish, faculty members of the Department of Computer Science and IT, IIS (deemed to be university), Jaipur, for their earnest involvement and commitment towards making this endeavour a success. We also thank our former-colleagues, Dr. Geetika Vyas and Ms. Joshita Sharma, for their preliminary contribution.

We also acknowledge the support provided by all other colleagues from IIS, (deemed to be university) Jaipur, for proficient management of the conference.

Finally, we are grateful to one and all, who have contributed directly or indirectly in making FICR-TEAS 2022 successful.

All this support came to its purpose with the blessings of the Almighty. We are indebted to him for showing us the right path throughout.

# Contents

<b>A Novel Approach for Detecting Online Malware Detection LSTM RNN and GRU Based Recurrent Neural Network in Cloud Environment .....</b>	1
M. Prabhavathy, S. Uma Maheswari, R. Saveeth, S. Saranya Rubini, and B. Surendiran	
<b>Analytical Study of Augmented Reality Simulation System Application on Paintings .....</b>	11
Priyanka Dewatwal and Nirupama Singh	
<b>A New Model SATV for Sentiment Analysis of Hinglish Sentences .....</b>	17
Shilpa Singh and Astha Pareek	
<b>Characterization of Molecular Dynamic Trajectory Using K-means Clustering .....</b>	25
Rajneet Kaur Bijral, Jatinder Manhas, and Vinod Sharma	
<b>Data Warehouse Modernization Using Document-Oriented ETL Framework for Real Time Analytics .....</b>	33
Monika Patel and Dhiren B. Patel	
<b>Opinion Mining for Breast Cancer Disease Using Apriori and K-Modes Clustering Algorithm .....</b>	43
S. Balakrishnan, K. Suresh Kumar, M. Balasubramanian, and Y. Rajkumar	
<b>Deep Learning Approach Based on ADASYN for Detection of Web Attacks in the CICIDS2017 Dataset .....</b>	53
Kuljeet Singh, Amit Mahajan, and Vibhakar Mansotra	
<b>FUNDUS and OCT Image Classification Using DL Techniques .....</b>	63
Akanksha Bali and Vibhakar Mansotra	

<b>A Secure Framework for the Deployment of Microservices Using Cloud Container Technology .....</b>	<b>77</b>
Salini Suresh, N. Ramachandran, M. Hanumanthappa, K. Ravikumar, and Anubha Jain	
<b>Time Series Analytics—A Fuel to Data Science in Forecasting Analytics for Time-Stamped Data .....</b>	<b>87</b>
Sunitha, Salini Suresh, Anubha Jain, and V. Suneeta	
<b>Sparse Autoencoder and Deep Learning Based Framework for Multi-label Classification of Chronic Diseases .....</b>	<b>97</b>
Atul Kumar Ramotra, Amit Mahajan, and Vibhakar Mansotra	
<b>A Comparative Study of Different Machine Learning Based Feature Extraction Techniques in Bacterial Image Classification .....</b>	<b>103</b>
Shallu Kotwal, Priya Rani, Tasleem Arif, and Jatinder Manhas	
<b>Cloud Removal and Satellite Image Reconstruction Using Deep Learning Based Image Inpainting Approaches .....</b>	<b>113</b>
Jaya Saxena, Anubha Jain, P. Radha Krishna, and Rajashree V. Bothale	
<b>A Novel Cipher Technique Using Substitution and Transposition Methods .....</b>	<b>123</b>
Ritu, Niram, Ekta Narwal, and Sumeet Gill	
<b>Edge Computing: State of Art with Current Challenges and Future Opportunities .....</b>	<b>131</b>
Vijay Singh Rathore, Vijeta Kumawat, B. Umamaheswari, and Priyanka Mitra	
<b>Deep Learning and Machine Learning Approaches for the Classification of Personality Traits: Review .....</b>	<b>139</b>
Jayshri Patil and Jikitsha Sheth	
<b>Two-Level Priority Task Scheduling Algorithm for Real-Time IoT Based Storage Condition Assessment System .....</b>	<b>147</b>
Saureng Kumar and S. C. Sharma	
<b>Intrusion Detection of Internet of Things Botnet Attacks Using Data Mining Technique .....</b>	<b>159</b>
S. Kavitha, M. Hanumanthappa, and B. N. Kalavathi	
<b>Code Tamper-Proofing Using Return Oriented Programming in IoT Devices .....</b>	<b>167</b>
Rajesh Shrivastava, Simar Preet Singh, and Mohammad Kamrul Hasan	
<b>Conversion of Punjabi Sign Language Using Animation .....</b>	<b>175</b>
Retinder Kaur and Suresh Kaswan	
<b>Deep Learning Based Classification of Microscopic Fungal Images .....</b>	<b>187</b>
Amit Sharma, Ajay Lakhnotra, Jatinder Manhas, and Devanand Padha	

Contents	xiii
<b>Application of Artificial Neural Network and Multiple Linear Regression in Predicting Consumer Attitude Towards Use of Mobile Wallet .....</b>	197
Anju Singh and Ruchi Jain	
<b>Securing Unauthorized Access to Cloud Data Storage .....</b>	205
Renu Devi, Ekta Narwal, and Sumeet Gill	
<b>A Survey on Water Usage Optimization Using Artificial Intelligence in Agricultural Sector .....</b>	215
N. Bharathi and M. Hanumanthappa	
<b>The Role of Artificial Neural Network in Word Sense Disambiguation (WSD)—A Survey .....</b>	221
H. R. Roopa and S. Panneer Arockiaraj	
<b>Covid-19 Challenges for Digital Education and Critical Analysis .....</b>	229
Puninder Kaur, Mehak Giri, Rajwinder Kaur, Avinash Sharma, and Manik Bhan	
<b>Cloud Computing: Concepts and Techniques with Service Provider Analysis .....</b>	237
Himanshu Barhaiya and Nilesh Jain	
<b>Analyzing Frequency Rhythms in Posed and Spontaneous Emotions Using Time-Intensity of EEG Signals .....</b>	251
Ritesh Joshi, Pratima Sen, and Maya Ingle	
<b>Improving Classification Performance Using the Semi-pivoted QR Approximation Algorithm .....</b>	263
Marina Popolizio, Alberto Amato, Vincenzo Piuri, and Vincenzo Di Lecce	
<b>Comparative Study on Video Based Face Recognition Methods .....</b>	273
Heta S. Desai and Atul M. Gonsai	
<b>An Innovative Framework Based Algorithmic Approach for Object Detection .....</b>	283
Anshika Jain and Maya Ingle	
<b>Smart Bus Concept of COVID-19 Prevention Using FaceBio Keys .....</b>	289
Dhara Rathore, Akash Saxena, and Gaurav Kumar Das	
<b>Data Integration Solutions—A Key to the Cloud Migration .....</b>	299
Ritul Bakshi and Vijay Singh Rathore	
<b>A Proposed Structure, Primarily Based Totally on Cloud Computing for Enhancing Electronic Commerce Applications .....</b>	307
Neha Tiwari and Navneet Sharma	

<b>Deep Learning and Internet of Things (IOT) Based Irrigation System for Cultivation of Paddy Crop .....</b>	319
S. Sasikumar, D. Naveen Raju, P. V. Gopirajan, K. Sureshkumar, and R. Pradeep	
<b>Review of Machine Learning Techniques for Analysis of Medical Data Sets .....</b>	329
Deepali Bhende and Gopal Sakarkar	
<b>Incorporating Contextual Information in Prediction Based Word Embedding Models .....</b>	341
K. Vimal Kumar, Shruti Ahuja, Skandha Choudary, and Pritee Parwekar	
<b>Distributed Cloud Data Center for Efficient Services in Logistics Management .....</b>	351
Shivani Dubey, Vikas Singhal, Rajanish Jain, and Amit Agrawal	
<b>Mining Plants Features for Disease Detection Tensor Flow: A Boon to Agriculture .....</b>	361
Saksham Goyal, Madhulika Bhatia, Urvashi, and Praveen Kumar	
<b>Data Visualisation Using Self-organising Maps .....</b>	373
Madhulika Bhatia, Saru, Preeti Manami, Pratham Vats, and Praveen Kumar	
<b>SQL Injection Attacks, Detection Techniques on Web Application Databases .....</b>	387
Navdeep Singh and Preeti Tiwari	
<b>Deep Learning Based Brain Tumor Segmentation: Recent Updates .....</b>	395
Rita B. Patil, Nirupama Ansingkar, and Prapti D. Deshmukh	
<b>Mobile Applications for Smart Agriculture in India: “An Analytical Approach” .....</b>	407
Madhu Gaur and Preeti Tiwari	
<b>Offline Handwritten Character Recognition of Gujarati Characters Using Convolutional Neural Network .....</b>	419
Pinal J. Patel, S. Kheni Rukshmani, Urmila Patel, Darshil P. Patel, Kinjal N. Patel, and Kinjal A. Patel	
<b>Forecasting Solar Energy on Time Frame: A Review .....</b>	427
Ashok S. Sangle and Prapti D. Deshmukh	
<b>Framework for Faster Index Search .....</b>	439
Chattar Singh and Vijay Gupta	
<b>Image Classification by Optimized Convolution Neural Networks .....</b>	447
Eva Tuba, Ira Tuba, Romana Capor Hrosik, Adis Alihodzic, and Milan Tuba	

Contents	xv
<b>Review: Benefits of Mining Fashion Data .....</b>	455
Monika Sharma and Priyanka Verma	
<b>Current State of Engineering Education in India: Student Teacher Scenario in Online Mode .....</b>	465
Aniruddh Sharan Mangla, Madhulika Bhatia, Shaveta Bhatia, and Praveen Kumar	
<b>Estimating Blockchain Using Time-Series Forecasting ARIMA .....</b>	477
Vaibhav Bhatnagar and Bhoomika Batra	
<b>Analyzing Online Reviews of Customers Using Machine Learning Techniques .....</b>	485
Anandhavalli Muniasamy and Roheet Bhatnagar	
<b>Assessing Between-Group Differences in Implementation of Lean Six Sigma Management Conception .....</b>	495
Slawomir Switek	
<b>Boosting Blockchain Mechanism Using Cryptographic Algorithm in WSN .....</b>	509
P. Anitha Rajakumari and Pritee Parwekar	
<b>Virtual Try-On Using Augmented Reality .....</b>	519
Aryan Khari, Muskan Gupta, Krishna Mohan, and Saru Dhir	
<b>Fractures Classification and Detection Using DIP and ML .....</b>	529
Nishit Dhamija and Saru Dhir	
<b>Sentiment Analysis in Social Media Website .....</b>	539
Aditya Rao, Mehak Khanna, Priya Sinha, and Saru Dhir	
<b>Artificial Intelligence Based Techniques for Underwater Image Restoration and Enhancement .....</b>	549
Yogesh Kumar Gupta and Khushboo Saxena	
<b>A Rapid Computing Technology on Profound Computing Era with Quantum Computing .....</b>	559
Ramnarayan, Minakshi Memoria, Anuj Kumar, and Sunil Ghildiyal	
<b>Covid-19: Comparison of Time Series Forecasting Models and Hybrid ARIMA-ANN .....</b>	567
N. Hema Priya, S. M. Adithya Harish, N. Ravi Subramanian, and B. Surendiran	
<b>YOLOV5 and Morphological Hat Transform Based Traffic Sign Recognition .....</b>	579
Ponduri Vasanthi and Laavanya Mohan	

<b>Optimizer Based Performance Study of Block Wise Fine-Tuned (BWFT) AlexNet Architecture .....</b>	591
Simran Kaur Hora and Maya Ingle	
<b>Comparative Study on Image Feature Extraction Techniques .....</b>	597
Yogita Bhajgotar and Vaibhav Gandhi	
<b>Cyber Security Threat, it's Risks and Susceptibilities, in the Global Perspective .....</b>	607
Radhika Mohan Gupta, Priyanka Mathur, and Ruchi Nanda	
<b>Detecting Cheating in Online Assessments Using Face Recognition and Head Pose Estimation .....</b>	615
Shaik Mohissin Sultana, M. Kameshwara Rao, and S. G. Santhi	
<b>Opinion System of Real-Time Emotions from Human Interaction Incorporated in Class Using Deep Learning .....</b>	623
M. Bindhu, S. Asha, and M. Praveen	
<b>A Study on Automatic Detection of Alzheimer's Disease Using Multimodalities .....</b>	631
Ag. Noorul Julaiha and R. Priyatharshini	
<b>IoT Enabled Smart Tourism (IoTeST): Tourism Service Dimensions .....</b>	643
Shivani Sharma and O. P. Rishi	
<b>An Innovative Approach Towards Designing Efficient K-Nearest Neighbour Algorithm .....</b>	655
D. Abhyankar and Maya Ingle	
<b>Latest Trends in Multi-modality Medical Image Fusion: A Generic Review .....</b>	663
Kapil Joshi, Mohit Kumar, Amrendra Tripathi, Anuj Kumar, Jitender Sehgal, and Archana Barthwal	
<b>Comparison Study of Classification Techniques for Predicting Performance of Students Using Weka Environment .....</b>	673
Shivani Sharma and Monika Rathore	
<b>Multimodal Autism Detection in Children Through Facial and Speech Emotion .....</b>	683
Monalin Pal and P. Rubini	
<b>A Contemporary Learning on Smart Devices for Smart Metropolis .....</b>	693
Sonali Vyas, Shaurya Gupta, Sourabh Singh Verma, and Vinod Kumar Shukla	
<b>Malpractice Detection in Online Assessments Using Eye Gaze Tracking and Object Detection .....</b>	701
B. Nagaraju Yadav and M. Kameshwara Rao	

Contents	xvii
<b>Big Data f5 Load Balancer with ChatBots Framework .....</b>	<b>709</b>
Kapil Joshi, Mohit Kumar, Minakshi Memoria, Pooja Bhardwaj, Gunjan Chhabra, and Dev Baloni	
<b>Keylogger Awareness and Use in Cyber Forensics .....</b>	<b>719</b>
Anuj Kumar, Kamlesh Kumar Dubey, Himanshu Gupta, Sumita lamba, Minakshi Memoria, and Kapil Joshi	
<b>Author Index .....</b>	<b>727</b>

## About the Editors

**Prof. (Dr.) Vijay Singh Rathore** obtained Ph.D. (Computer Science), MCA, M.Tech. (CS), MBA, ICAD-USA, ICDA-USA, and Diplomas in French & German. Presently, he is working as Professor of CS, Director—International at IIS (deemed to be university), Jaipur (India), Secretary, ACM Jaipur Chapter, and Past Chairman, CSI Jaipur Chapter. He is having teaching experience of 20+ years, three patents published, Ph.D. supervised (awarded: 18, under supervision: 9), published 94 research papers, and ten books, associated as Editor and Reviewer in some reputed journals, and received 20+ national and international awards of repute. His core research areas include Internet security, cloud computing, big data, and IoT. He has organized and participated 25+ national and international conferences of repute. His foreign visits for various academic activities (Delegate/Invited/Keynote Speaker/Session Chair/Guest) include USA, UK, Canada, France, Netherlands, Singapore, Thailand, Vietnam, Nepal, etc. He had been Member of Indian Higher Education Delegation and visited 20+ leading universities at Canada (May, 2019), UK (July, 2017), and USA (Aug., 2018) supported by AICTE and GR foundation. Other delegation visits include University of Amsterdam (2016), Nanyang Technological University, Singapore (2018), University of Lincoln, QMUL, Brunel University, and Oxford Brookes University (2020) for discussions on academic and research collaborations. He is Active Academician and always welcoming and forming innovative ideas in the various dimensions of education and research for the better development of the society.

**Prof. Subhash Chander Sharma** (<https://www.iitr.ac.in/~DPT/scs60fpt>) received M.Sc. (Electronics, IITR), M.Tech. (Electronics and Communication Engineering, IITR), and Ph.D. (Electronics and Computer Engineering, IITR) in 1981, 1983, and 1992, respectively, from IIT Roorkee (erstwhile University of Roorkee). He started his career as R&D Engineer in 1983 and then joined the teaching profession in January 1984 at IIT Roorkee and continuing till date. He has more than 35 years of teaching and research experience at IIT Roorkee. He has published over three hundred fifty research papers in national and international journals (152)/conferences (150) and supervised more than 30 projects/dissertations of PG students. He has supervised 20

Ph.D. in the area of computer networking, wireless network, computer communication, cloud and its security, mobile computing, and continuing supervising Ph.D. in the same area. He has successfully completed several major research projects funded by various government agencies like AICTE, CSIR, UGC, MHRD, DST, DRDO, and many minor research projects related to communication and SAW filter design sponsored by the Government of India. IIT Roorkee has awarded him the Khosla annual research prize with the best research paper. His many research papers have been awarded by national and international committees and journals. He has worked as Research Scientist at FMH, Munchen, Germany, and visited many countries (UK, France, Germany, Italy, Switzerland, Canada, UAE, Thailand, Netherlands, etc.) related to research work. He has chaired sessions of international conferences and delivered invited talks at various forums. He is Active Reviewer of the IEEE Journals and Editor of various reputed international and national journals. He is Honorary Member of NSBE, ISOC, and IAENG, USA. He has also worked as Group Leader of Electronics and Instrumentation Engineering Department of BITS Pilani, Dubai Campus, from August 2003 to August 2005. Presently, he is continuing as Professor at IIT Roorkee, Saharanpur Campus.

**Joao Manuel R. S. Tavares** graduated in Mechanical Engineering at the Universidade do Porto, Portugal, in 1992. He also earned his M.Sc. degree and Ph.D. degree in Electrical and Computer Engineering from the Universidade do Porto in 1995 and 2001 and attained his Habilitation in Mechanical Engineering in 2015. He is Senior Researcher at the Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial (INEGI) and Associate Professor at the Department of Mechanical Engineering (DEMec) of the Faculdade de Engenharia da Universidade do Porto (FEUP). He is Co-editor of more than 60 books and Co-author of more than 50 chapters, 650 articles in international and national journals and conferences, and three international and three national patents. He has been Committee Member of several international and national journals and conferences, is Co-founder and Co-editor of the book series *Lecture Notes in Computational Vision and Biomechanics* published by Springer, Founder and Editor-in-Chief of the journal *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* published by Taylor & Francis, Editor-in-Chief of the journal *Computer Methods in Biomechanics and Biomedical Engineering* published by Taylor & Francis, and Co-founder and Co-chair of the international conference series: CompIMAGE, ECCOMAS VipIMAGE, ICCEBS, and BioDental. Additionally, he has been (Co-)supervisor of several M.Sc. and Ph.D. thesis and Supervisor of several postdoc projects and has participated in many scientific projects both as Researcher and as Scientific Coordinator. His main research areas include computational vision, medical imaging, computational mechanics, scientific visualization, human–computer interaction, and new product development.

**Catarina Moreira** is Computer Scientist and passionate about decision-making technologies that can empower human decisions. She is Lecturer in Information Systems at the Faculty of Science, Queensland University of Technology, where

she holds the role of Deputy High Degree Research Academic Lead. Currently, she is Associate Editor for the Springer Nature journal BMC Bioinformatics, section “Machine learning and artificial intelligence”. She is Pioneer in the development of quantum-like probabilistic graphical models for decision-making to empower human decision-making. Her work also comprises machine learning/deep learning models and innovative human interactive probabilistic models for explainable AI with the goal of understanding how and why machines make decisions. One of her aspirations is to support women in research and in STEM (Science, Technology, Engineering, and Mathematics) with the aim of contributing to a society that promotes equal opportunities across all genders to engage in STEM-related areas.

**Dr. B. Surendiran** is working as Assistant Professor, Department of CSE, NIT Puducherry, Karaikal. He has more than 15 years of teaching and research experience. He had completed his Ph.D. in the year 2012 from NIT Trichy. His research interests are medical image processing, machine learning, dimensionality reduction, etc. He had 40 international journal and conference publications to his credit. He has delivered 75+ invited lectures, reviewed 350+ research papers, and guided ten PG/Ph.D. scholars. He has received Best Reviewer Award in 2019 by publons.com, Web of Science, and also Outstanding Reviewer Award by Biomedical and Pharmacology Journal in the year 2020. He is Active Reviewer for IEEE sensors, IET image processing, and various SCI/Scopus Journals.

# A Novel Approach for Detecting Online Malware Detection LSTM RNN and GRU Based Recurrent Neural Network in Cloud Environment



**M. Prabhavathy, S. Uma Maheswari, R. Saveeth, S. Saranya Rubini, and B. Surendiran**

**Abstract** Now a day's most of the organizations depend on cloud infrastructure for application, storage and real time access perspective. Cloud based application save as a backbone to organization in terms of maintainability, scalability and management underlying infrastructure. To meet the demand of such organization cloud service providers provide Infra-Structure as a Service (IaaS) through Amazon EC2, Microsoft and Azure Virtual Machine. Because of higher utilization of cloud platform, it becomes more targeted to attackers. In security view of IaaS, malware become the most dangerous threat to IaaS Infrastructure. In proposed method, two RNN architectures of deep learning are considered for detecting malware in cloud virtual machine such as LSTM RNNs (Long Short Term RNN) and GRU (Gated Recurrent Unit RNN). Behavioral features of CPU, disk utilization and memory are learnt. Running applications in real time environment in online cloud platform with no restriction and capture all behaviors of normal and benign applications. This model achieves higher detection rate of 99% over 42,480 datasets. Based on the ordering of input data, the performance varied.

**Keywords** Deep learning · LSTM RNN · GRU · Cloud computing · IaaS · Cloud service providers

---

M. Prabhavathy (✉) · R. Saveeth · S. Saranya Rubini

Department of CSE, Coimbatore Institute of Technology, Coimbatore, India

e-mail: [prabhavathy.phd@gmail.com](mailto:prabhavathy.phd@gmail.com)

R. Saveeth

e-mail: [saveeth.r@cit.edu.in](mailto:saveeth.r@cit.edu.in)

S. Saranya Rubini

e-mail: [ssaranya@cit.edu.in](mailto:ssaranya@cit.edu.in)

S. Uma Maheswari

Department of ECE, Coimbatore Institute of Technology, Coimbatore, India

B. Surendiran

Department of CSE, National Institute of Technology, Puducherry, India

e-mail: [surendiran@nitpj.ac.in](mailto:surendiran@nitpj.ac.in)

## 1 Introduction

Cloud infrastructure is most secured environment that need now days. Because of more popularity more chance of data stealing. Cloud provides many features in addition to that it is available to both private and public sectors based on pay-per usage concept. Organization having many branches they no need to carry their data individually to transfer, instead started using IaaS [1] Service. In IaaS service, client can make use of virtual machine using computational resources and satisfy their needs. If attackers gain access through IaaS, entire organization data losses occur. So securing cloud infrastructure is more complicated [2]. In cloud IaaS platform, malicious apps are executed to attack the VMs. Some of the challenges are automated configuration tool, number of clients increasing every year, inability to secure the resources acquired. When recreating a VM, configuration vulnerability that using a script is spawned automatically may be prone to attack. Further any more VM spawned it causes same problem. When deploying configuration script for more number of times, cloud malware easily propagate to each server and detecting such malware is complicated task for cloud service providers.

To overcome many challenges, many malware detection methods available and categorized as static analysis [3], dynamic analysis [4], hybrid analysis [5] and memory analysis. Static analysis method works on signature based anomaly detection method that results in identifying static features. Dynamic analysis method based on running an executable file and identifying its features and behaviors. Both methods have advantages and disadvantages. Static malware approach failed to detect zero day malware and some malware changes its behavior (polymorphic) dynamically so it's difficult to identify features. However malware detect the sandboxes and stop the malicious actions to evade detection. Both methods failed to detect executable before they get into host. To overcome the drawbacks in both methods, here online malware detection method is used to differentiate benign from malicious application. Identifying malware on on-demand basis is essential in cloud environment, so our approach is a reliable way for all methods of malware detection. In our research, Recurrent Neural Networks (RNNs) with online malware detection method is used for identify the system features of virtual machine in IaaS infrastructure. Many VMs are running in cloud, each contain fixed set of processes. When executing a process, many new processes are creating and deleting it. Ex: database process, web application process, web server process. Some sequence order maintained in executing a process, if malicious process executing its code, order varies. When there is change in sequence, using RNN identify the malicious application from benign is easy.

During analysis, showed that malware changing their process name to system process name. Malware attach itself to legitimate process and behave like legitimate process. To overcome the limitations of shallow CNN model, we proposed a new approach RNN based deep learning method. The work in our paper is spitted as follows: Proposed a novel approach for detecting process system features of cloud malware using RNNs. Demonstrated the set of unique processes running in virtual machine as sequence of process system features. RNNs based method identifying all

types of process sequence. Provided LSTMRNN and GRU for analyzing evaluation metrics and also detection time and training time. Different types of input representations are also analyzed. Irrespective of the order of system features, LSTMRNN and GRU achieves high performance with range 0.1% to 0.3%.

## 2 Literature Survey

Dynamic malware detection and online malware detection are two methods based on behavioral based ML methods for detecting malware. There exists a major difference in two methods: executable files are running in sandbox and capturing its behavior in dynamic malware detection method, where in online malware detection method, real time data's are analyzed and labeled as malicious if identified.

### 2.1 Dynamic Malware Detection Method

Many traditional machine learning algorithms are identified for malware detection. In [6] traditional machine learning algorithms are used which extracted system calls. Ex SVM, Decision tree, KNN, Naïve Bayes.

In traditional machine learning algorithms neural network concepts are included. In [7] author identified pattern using sequence mining technique and it utilizes a Nearest Neighbor classifier to identify unknown malware. In [8] author used system performance metrics and ML concept. More sophisticated methods are needed than traditional method because traditional method takes more time to pre-processing and extracting features. Instead of wasting time in extracting features in ML, moving towards deep learning methods. Dynamic method specifically works on host-based system instead of dependent on cloud platform and limited internet access. To overcome the above limitations, online malware detection methods are used.

### 2.2 Online Malware Detection Method

In paper [9], the author used traditional machine learning method along with memory features are considered. In paper [10], author used traditional machine learning algorithm such KNN, Naïve Bayes for malware detection and used system calls for intrusion detection. To overcome accuracy limitations in cloud, the author [11] proposed system calls based anomaly detection for VM in cloud environment. They used decision tree and Naïve Bayes method. Many features are extracted from hypervisor by cloud service providers because they do not required to see inside VM. Performance

metrics are also considered and using SVM classifier. But it identifies only highly utilized resources, low resource utilization remain unidentified.

### 3 Proposed Work

#### 3.1 LSTM RNNs Model

The mostly used model in deep learning is recurrent neural network which best suits for time series prediction problem. Even though RNN have disadvantages such as short term memory loss that is forgetting previous information when large number of input passed and during backpropagations vanishing gradient value. LSTM based RNNs model solved the above limitations. LSTM RNNs model consist of input, output and forget. It keeps the necessary information to cell. Using sigmoid function, forget unit decides which data should be forgotten. Using point wise multiplication operation such as “sigmoid” and “tanh” input unit control current information to current cell. The sigmoid activation function ranges between 0 and 1. If the value closer to 0, then data may be discarded and it is considered as lesser important. If the value is nearer to 1, then data may be retained and considered as more important. To control the output tanh value should be –1 and 1. The output unit finally decides which data to be passed to next hidden layer. By considering the entire factors LSTM RNN model is a reliable model without losing data.

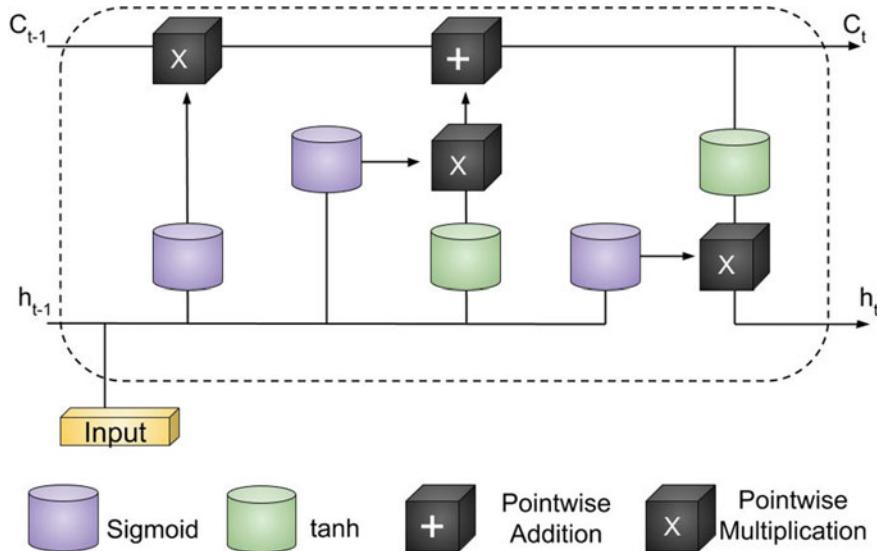
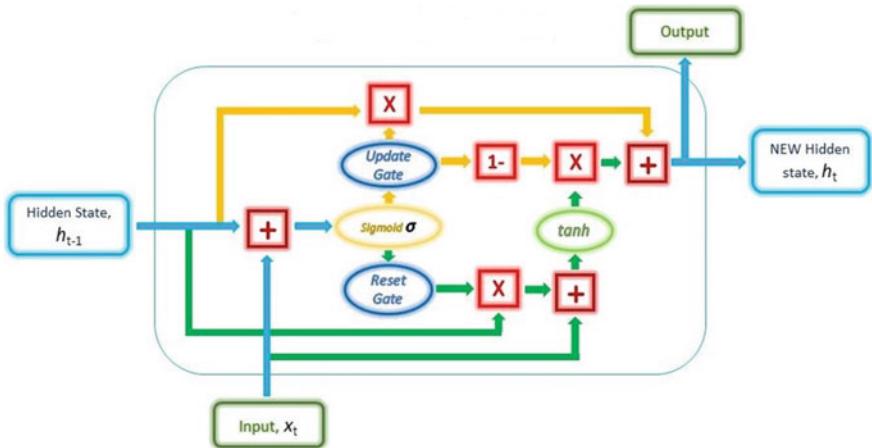
#### 3.2 GRU Model

Figure 2 shows the GRU RNN architecture. It contains only two gates, such as Update gate and Reset gate. It works similar to LSTM, GRU gates are trained to select only the necessary information and forget the irrelevant information. The values associated between 0 and 1 for the gates, which will then be multiplied with input data and hidden state. It will return zero, when values 0 in gate vectors in input or hidden state is irrelevant information. On the other side, when values 1 in gate vectors in input or hidden state is useful information (Fig. 1).

## 4 Experimental Analysis and Result

#### 4.1 LSTM RNN and GRU Models

In our proposed methodology LSTM RNN model is the first model which contains 8 layers architecture. The first three layers consist of the following units such 256,

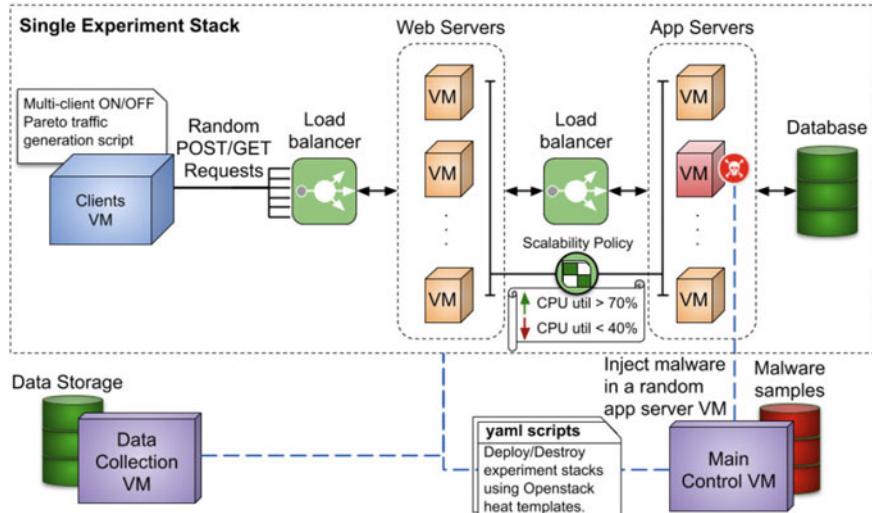
**Fig. 1** LSTM architecture**Fig. 2** Architecture of a GRU unit

128 and 64 respectively. To overcome over fitting LSTM-RNN followed by 10% of dropout layer. Softmax activation function present in final output layer. To check whether it is malicious or benign, binary classification method is used. GRU is second model and this model contains 4 directional LSTM unit and neurons such as 512, 256, 128, 64 respectively and 10% dropout layer is followed. Softmax activation function is present in two output layer. mpk denotes malicious sequence. 125 experiments are analyzed with dataset are trained, validated and tested, where 60% of data's are

used for training, 20% of data's are used for validation, 20% of data's are used for testing purpose. For obtaining hyper parameter optimization grid search method is used to get better model and sizes are 16, 32, 64 and learning rates are 0.0001, 0.0001, 0.00001 etc.

## 5 Experimental Analysis

Cloud deployments are represented in Fig. 3. To deploy that we use 3-tier architecture consist of the web server, the application server and the database server. A load balancer is available at front to transfer the incoming client request to appropriate web server. In turn web servers are connected to application server to distribute the load equally to application server. At the end application server are connected to database server. Auto scaling option is implemented, it automatically starts working when CPU usage of VMs exceeded 70% and fall below 40%.The experiment is conducted for totally one hour. First half an hour for benign application, no malwares are identified. A malware file is injected at random time between 30 to 40 min. After 40 s malicious samples are taken for analysis. Collecting data samples ever 10 sends. Totally 360 samples taken for experiments. At the end 42,480 samples are taken for analysis.

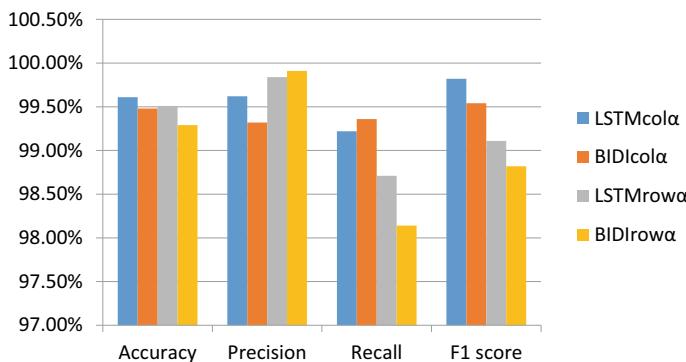


**Fig. 3** Experiments deployment

## 5.1 Results

For our experimental analysis, 118 experiments are taken and training 60% data, validation 20%, testing purpose 20% data are used. For ensuring zero-day malware, datasets are divided into no of experiments. (i.e. 68% training, 25 validation, 25 testing). The data samples (25) collected from unseen malware are unseen to RNN model. The training datasets are used for training RNN model, validation dataset are used to get better model using parameter such as batch-size, learning rate and to measure the detection ratio of LSTM RNN and GRU the testing datasets are used. The models are trained to ensure validity of results (i.e. LSTM RNN $\alpha$ , LSTM RNN $\beta$ , GRU $\alpha$  and GRU $\beta$ ). To get dynamism, orders are get varied (i.e.  $\alpha$  and  $\beta$  are modified). The models trained for 40 periods but for loss there is no decrease for further periods (Fig. 4 and Table 1).

Table 2 represents epochs where the highest validation accuracy is achieved. In general, GRU model converged relatively faster than LSTM model. The GRU  $\gamma$ , GRU  $\delta$ , GRUrow  $\gamma$ , GRU col  $\gamma$ , GRUrow  $\delta$ , GRUcol  $\delta$  converged fastly with respect to the values 2850, 2889, 3002, 2884, 4520, 2765. But the LSTM model LSTM  $\gamma$ , LSTM  $\delta$ , LSTMrow  $\gamma$ , LSTMcol  $\gamma$ , LSTMrow  $\delta$ , LSTMcol  $\delta$  take more time to converge and their values respectively are 4189, 5244, 6115, 3189, 6110, 3100 s. The results show LSTM model take more time to converge because many backward layers are added.



**Fig. 4** LSTM RNN and BRNN ( $\alpha$  and  $\beta$ ) models results

**Table 1** LSTM RNN and GRU ( $\alpha$  and  $\beta$ ) models results in detail

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
LSTMcol $\gamma$	99.61	99.62	99.22	99.82
GRUcol $\gamma$	99.48	99.32	99.36	99.54
LSTMrow $\gamma$	99.51	99.84	98.71	99.11
GRUrow $\gamma$	99.29	99.91	98.14	98.82

**Table 2** Optimal LSTMRNN and GRU model training time

Model	Validation accuracy (%)	Epoch reached	Time elapsed (s)	Loss (%)
LSTM $\gamma$	99.62	10	4189	3.01
GRU $\gamma$	99.95	27	2850	2.75
LSTM $\delta$	99.65	37	5244	1.54
GRU $\delta$	99.98	32	2889	0.73
LSTMrow $\gamma$	99.77	40	6115	2.30
GRUrow $\gamma$	99.64	40	3002	1.70
LSTMcol $\gamma$	99.79	38	3189	2.93
GRU col $\gamma$	99.96	19	2884	2.50
LSTMrow $\delta$	99.65	32	6110	3.11
GRUrow $\delta$	99.94	29	4520	2.65
LSTMcol $\delta$	99.62	39	3100	1.64
GRUcol $\delta$	99.93	32	2765	0.83

## 6 Conclusion

In this work, a method to use the LSTMRNN is proposed and GRU models to detect online malware with unique process characteristics. The proposed model works best on the test data set. However, the GRU model requires less time than the LSTM model to achieve better performance. Change the input representation of the random sort row and the random sort column to identify unique characteristics and processes. When sorting is changed, performance is not affected, but it does affect model training time. A sample of the dataset was used, but in the future, thousands of samples will be used and also try to identify what kinds of malware families it belongs to. In the proposed method, two deep learning RNN architectures are considered to detect malware in cloud virtual machines such as LSTM RNN (Long Short Term RNN) and GRU (Two-way RNN). Behavioral characteristics of processor, disk usage and memory are learned. Run applications in a real-time environment on an unrestricted online cloud platform and capture all normal and benign application behavior. This model achieves a 99% higher detection rate on 42,480 data sets. Based on the sorting of the input data, performance varied.

## References

1. S. Namasudra, P. Roy, B. Balusamy, Cloud computing: fundamentals and research issues, in *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)* (2017), pp. 7–12. <https://doi.org/10.1109/ICRTCCM.2017.49>
2. X. Liu, C. Xia, T. Wang, L. Zhong, CloudSec: A novel approach to verifying security conformance at the bottom of the cloud, in *2017 IEEE International Congress on Big Data (BigData Congress)* (2017), pp. 569–576

3. M. Ijaz, M.H. Durad, M. Ismail, Static and dynamic malware analysis using machine learning, in *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)* (2019), pp. 687–691. <https://doi.org/10.1109/IBCAST.2019.8667136>
4. M. Egele, T. Scholte, E. Kirda, C. Kruegel, A survey on automated dynamic malware-analysis techniques and tools. *ACM Comput. Surv.* **44**(2), 1–42 (2012)
5. Z. Bazrafshan, H. Hashemi, S.M.H. Fard, A. Hamzeh, A survey on heuristic malware detection techniques, in *2013 5th Conference on Information and Knowledge Technology (IKT)* (IEEE, 2013), pp. 113–120
6. I. Firdausi, C. Lim, A. Erwin, A.S. Nugroho, Analysis of machine learning techniques used in behavior-based malware detection, in *Proceeding of 2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies* (Dec 2010), pp. 201–203
7. Y. Fan, Y. Ye, L. Chen, Malicious sequential pattern mining for automatic malware detection. *Expert Syst. Appl.* **52**, 16–25 (2016)
8. S. Joshi, H. Upadhyay, L. Lagos, N.S. Akkipeddi, V. Guerra, Machine learning approach for malware detection using random forest classifier on process list data structure, in *Proceedings of the 2nd International Conference on Information System and Data Mining (ICISDM)* (2018), pp. 98–102
9. J. Zhang et al., Malware detection based on dynamic multi-feature using ensemble learning at hypervisor. *IEEE Glob. Commun. Conf. (GLOBECOM) 2018*, 1–6 (2018). <https://doi.org/10.1109/GLOCOM.2018.8648070>
10. Z.S. Malek, B. Trivedi, A. Shah, User behavior pattern-signature based intrusion detection, in *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)* (2020), pp. 549–552. <https://doi.org/10.1109/WorldS450073.2020.9210368>
11. Q. Guan, Z. Zhang, S. Fu, Ensemble of Bayesian predictors for autonomic failure management in cloud computing, in *2011 Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)* (2011), pp. 1–6. <https://doi.org/10.1109/ICCCN.2011.6006036>

# Analytical Study of Augmented Reality Simulation System Application on Paintings



Priyanka Dewatwal and Nirupama Singh

**Abstract** Augmented reality technology is a revolution of the 21st-century on the other side paintings are a traditional tool of expression. Representation of painting through technology is innovative, effective, and exciting to watch. The use of augmented reality is rising in every sector, and it has changed the perception of paintings in the last few years. As we know paintings are the most beautiful and famous way to express an artist's feelings and thoughts on canvas, and this time art and technology collaborative study gave an extraordinary and unique result. This research paper is about the augmented reality simulation system application on paintings and its communicating presentation. It is an interactive installation art due to the involvement of spectators. Augmented Reality paintings are demanding in the international art market, as artists accepted that augmented reality interferes in fine arts territory. This paper is about the process and analytical study of augmented reality simulations on paintings with conceptual thinking and its deliberation as a new media art in the international art market.

**Keywords** Augmented reality · Augmented reality simulation system · Technology · Painting · Fine arts · International art market

## 1 Introduction

Contemporary paintings diameter is increasing gradually and globally with the help of astonishing technologies. Contemporary artists are using augmented reality technology widely in the art sector. Augmented reality paintings are the invention by the technological artist where spectators can see moving objects or characters on the paintings. It can be a celebrating moment for not only a viewer but for a traditional painter also. Traditionally, artists have experimented with various mediums, surfaces, and the mixing of multiple mediums to create paintings, but in recent decades, artists have begun to use digital platforms to create unexpected artworks

---

P. Dewatwal (✉) · N. Singh  
IIS (Deemed to be University), Jaipur, India  
e-mail: [priyanka.dewatwal@gmail.com](mailto:priyanka.dewatwal@gmail.com)

such as digital paintings, installation art, digital interactive media, new media art, audio-video interactive art, and now augmented reality paintings. In order to produce augmented reality paintings, artists first create a painting on a 2D surface, then digitally transfer it using 3D software, and finally simulate both using augmented reality programmes.

## 2 Augment Reality Appearance in Paintings

Augmented reality is a powerful tool that is utilised in a variety of disciplines such as education, entertainment, medicine, aviation, and many others in today's world, but this application in painting is novel and revolutionary. In India, the emergence of augmented reality technology in artworks appears to have occurred between 2010 and 2014. Numerous traditional and technical artists in India developed an interest in new media art and participated as contemporary artists in many national and international art festivals. Animated paintings are a common term these days. The digital version of Indian miniature paintings developed by Google to highlight rich Indian culture and tradition is the best example of augmented reality painting [1]. Art galleries in Hyderabad, Mumbai, Delhi, and Bengaluru are supporting augmented reality artists and presenting their work around the world. They support and purchase the work of augmented reality artists. These paintings are more collaborative, appealing, and unique than other contemporary art forms, and they are significant as twenty-first-century media art (Fig. 1).

### 2.1 Related Works

Augmented reality is now a required component of new media art [2]. Several artists are using self-developed techniques and related animated art to create augmented reality paintings. During the pandemic, many art galleries used augmented reality techniques for online exhibitions to engage visitors, buyers, and art enthusiasts. People use augmented reality software to scan the painting and observe the animated figures and objects on their phones or tablets. The advent of augmented reality paintings, in which artists convey their conceptual ideas using fundamentals of fine arts, is due to the usage of artificial intelligence technology in new ways.

The creation of augmented reality paintings is a typical process in which artist ideas collide with technology as they transition from pencil and paper to AR. Artists use paint materials to paint on paper or canvas, then generate a digitally animated version to emulate with an AR application like "artviva." These artworks are a great example of people's positive and negative thinking levels. Moving visuals throughout the application can be more impactful than fixed images since they allow the viewer to interpret the painting according to their mental shape.



**Fig. 1** Indian miniature painting digital version

Examples of augmented reality contemporary artworks include “Priya’s Shakti,” “Alappuzha,” and “Juggernaut.” Priya’s Shakti is a painted-style conceptual comic series that concerns rape cases, gender-based violence, and sexual harassment situations that are on the rise in society. Priya, which means beloved, is a rape survivor, and Shakti, which means power to substance for a change, is the project’s theme [3]. The Blippar programme allows users to see augmented reality street paintings. Mahendra Parshuram Vartak, Niren Vashram Savaniya, and Sham Maruti Jadhav created these life-size paintings on the walls of Dharavi, Mumbai, India. The goddess Durga appears to be perched atop a tiger in an agitated mood after scanning the murals with an AR programme. Priya’s shakti website/application offers a digital replica of these murals (Fig. 2).

“Alappuzha” is an example of augmented reality art developed by Fabin Rasheed, an artificial intelligence, augmented reality, and virtual reality artist from India. His contributions to art and technology are well-known around the world. “Alappuzha” is an interactive live generative art piece based on the spectator’s conceptual observation during the show. Alappuzha is a southern Indian city with magnificent beaches and farmland. This work of art is a blend of changing locations, times of day, weather, and city seasons. Artist made 1024 different scene images of five key metropolitan places, each with its own time (morning, evening, night, afternoon), season, and time (morning, evening, night, afternoon) (summer, winter, rain, etc.). These illustrations are available for viewing on the artist’s website (Fig. 3) [4].

Sahej Rahal’s “Juggernaut” paintings are an example of animation simulations generated with an augmented reality application system [5]. According to the artist,

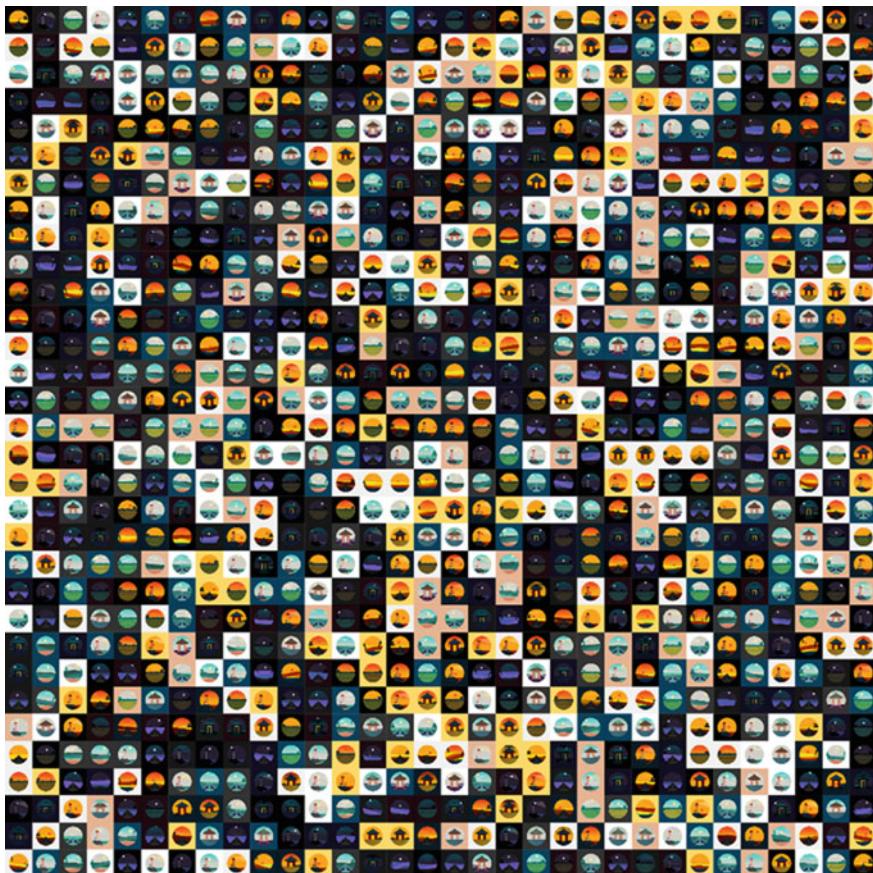


**Fig. 2** Priya Shakti wall mural

“Juggernaut” is corruption at the Jagannath Vishnu temple. Sahej painted paintings of numerous gods with hundreds of arms and animal heads, complicated shapes, disproportionate figures, and amazing appearances akin to superheroes [6]. While observing through an application, elements such as the tail, eyes, hands, and legs are visible moving. That is how alternative civilizations are represented. This artwork is a one-of-a-kind revolutionary work of contemporary art due to its powerful visualisation of shapes, colours, and forms, as well as its use of technology (Fig. 4) [7].

### 3 Conclusion

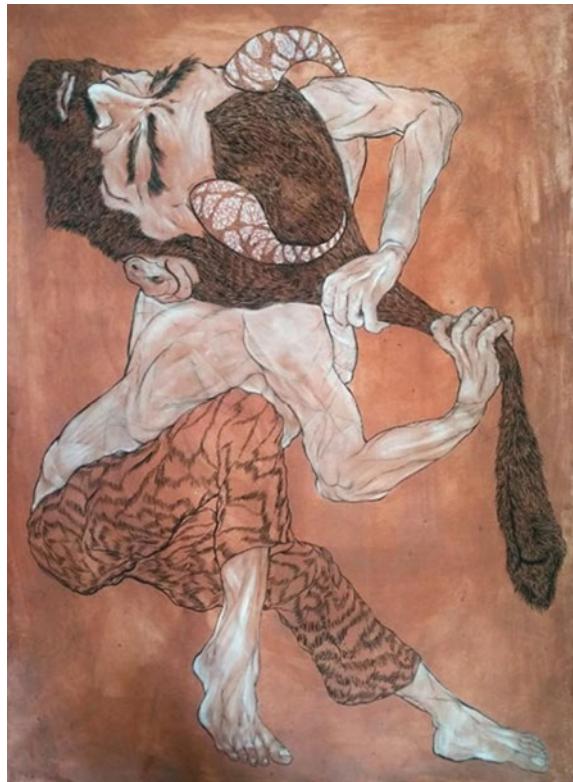
Digital NFTs and crypto art, a subset of augmented reality art, are becoming increasingly popular among young people and in the art industry. Art, animation, computer science, and artificial intelligence are all used to create this work. Both fine art and technology graduates are interested in learning how to use augmented reality to produce mind-blowing masterpieces; they represent the virtual art world’s future. Spectators are drawn to the fascinating artificial world, which encourages them to attend art events all over the world. The audience is awestruck as they enter the artificial world of fiction, truth, fantasy, psychology, good vs evil, and thinking manipulations. On social networking platforms like Instagram, Facebook, Twitter, and LinkedIn, the popularity of augmented reality painting may simply be calculated.



**Fig. 3** Variations of combined illustration into one single generative artwork by Fabin Rasheed

Traditional art education and technological knowledge are important components of creating great conceptual content. The popularity of augmented reality art could be a reason for it being the next component of the art in the future, and art academies will most likely include it as a part of visual arts in the near future.

**Fig. 4** Juggernaut by sahej rahal



## References

1. Buzz Staff, Google to Bring Indian Masterpieces to Life with AI, AR-Powered Virtual Art
2. Gallery for Miniatures, *News18*, 25 Oct 2020
3. <https://www.news18.com/news/buzz/googles-virtual-gallery-of-miniature-paintings-to-bring-indian-masterpieces-on-smartphones-using-ar-3005282.html>
4. Meet Three Leading AI-Based Artists in India, *Analytics India Magazine*, 13 Oct 2021. <https://analyticsindiamag.com/meet-three-leading-ai-based-artists-in-india/>
5. Rattapallax, Animated street art in Mumbai, India. *Medium*, 25 Aug 2014. <https://medium.com/@rattapallax/animated-street-art-in-mumbai-india-4aec932b2544>
6. Alappuzha. *Nurecas*. <https://nurecas.com/alappuzha>
7. “Art | Sahej Rahal’s ‘Juggernaut’ at Mumbai’s Chatterjee & Lal.” *YouTube*, 24 Feb 2019. <https://www.youtube.com/watch?v=G40hHIQH8uI>

# A New Model SATV for Sentiment Analysis of Hinglish Sentences



Shilpa Singh and Astha Pareek

**Abstract** With the increased use of the Internet and Social media platforms. It becomes very easy to know the opinion and thoughts of people through social media. Users may share their personal feeling in their local language. These tweets may be used to analyze the sentiments of people through Sentiment Analysis. In India, most of the people use Hinglish (Hindi words written in English) or Indian English language to share their feeling. So, in this paper, we have proposed a new model named SATV (Sentiment Analyzer Translating Vader) for Sentiment Analysis of Hinglish Language sentences. This model has been implemented and tested on Delhi University tweets taken from Twitter to examine the flow and working of the SATV model. Linear SVC and Naive Bayes Classifications have been applied for calculating the accuracy. Also, Precision, Recall, and F-measure have been calculated using a confusion matrix. Lastly, a comparison of the SATV model has been done with Vader Technique.

**Keywords** Hinglish sentiment analysis · SATV · NLTK · Naïve Bayes · Linear SVC · Python · Twitter · Vader

## 1 Introduction

From the last decade, with expeditiously increasing use of the internet, there has been an extreme rise in the sharing of thoughts on microblogging and blogging. The majority of people prefer Twitter.com to express their feelings on numerous topics like politics, sports, entertainment, education and etc. People like to prefer their local or national language to write tweets on Twitter [1]. In India, people use many regional languages to express their feelings on Twitter. In the same way, a large

---

S. Singh (✉) · A. Pareek  
Department of CS& IT, The IIS University, Jaipur, India  
e-mail: [shilpamca85@gmail.com](mailto:shilpamca85@gmail.com)

A. Pareek  
e-mail: [astha.pareek@iisuniv.ac.in](mailto:astha.pareek@iisuniv.ac.in)

group of people from India prefer Hinglish (Indian English) language to write tweets [2].

Several researches have been done on sentiment analysis of different languages like Tamil, Bengali, Spanish, etc. [3]. But less work has been done on Hinglish Sentiment Analysis. Sentiment Analysis of this Hinglish language is a challenging task because the words of these Hinglish sentences are not in fair Hindi or English language. For example:

Please college khol do.... humme depression ho rhe hai ghar me

In above example some words are in English like Please, college and depression. And some are in Hinglish language like khol, do, humme and so on.

Highly use of this type of language on social media not only brings new opportunities but also opens a new area of research as it has lots of challenges making it a prime area of research.

This paper is focused on proposing a new model SATV for doing Sentiment Analysis of Hinglish sentences. Delhi University tweets have been taken from Twitter using the hashtag “delhiuniversity”. This data is based on the liking and disliking of online classes held because of the COVID19 pandemic.

The remnant of the paper has been prepared in 5 sections. Section 2 gives a supervision of the literature review that represents the work done on Hinglish Sentiment Analysis. In Sect. 3, the polarity of each tweet has been calculated using Vader and our proposed model SATV then accuracy has been calculated. Section 4 shows the comparison between Vader and SATV model for accuracy, precision, recall, and Fmeasure. Finally, future research with the conclusion is reported in Sect. 5.

## 2 Literature Review

Numerous researches have been done on multilingual sentiment analysis. Author [4] proposed a new model for sentiment analysis of Hindi sentences. Author Kameshwar first created a seed list of the most frequently used word with polarity. Firstly words extracted from Tagging step of preprocessing of data have been matched with the words available in the seed list. If the match is found then polarity has been calculated according to the seed list. But, if the match does not found then the synonym of that word has been searched from the Hindi Dictionary manually created by the author. If a match is found in Dictionary then that synonym word with their polarity has been updated to the seed list. So, this seed list is updated every time a new word is found. Accuracy, precision, and recall have been calculated using threefold cross-validation with 64.05, 63.96, and 44.75 respectively. The author used the k fold cross-validation technique for finding accuracy but other techniques may also be used for finding accuracy.

Many authors like [5, 6] proposed a new technique by enhancing SentiWordNet for Sentiment Analysis of Hinglish Sentences. Author Pranaydeep used both Supervised and Unsupervised techniques with SeedDect for finding sentiments of Hinglish

sentences. While author Seshadri manually assigned polarity values to the words of Hinglish language to train data then whole data has been tested on the basis of this trained data. Author used a supervised system architecture that based on Bidirectional LSTM model for Hinglish sentiment analysis.

Arabic Sentiment Analysis has been done by the author Gamal [3]. In this paper author first, collect Arabic data from Twitter then preprocessing of data has been done. After that, various Machine learning Techniques like NB, SVM, and ME algorithms have been used for training and testing of data with the tenfold cross validation process. Author [7] done a literature review on Hinglish language sentiment analysis Hinglish language is the combination of English and Hindi i.e. in one sentence some words are in English and some words are in Hindi language. A comparative study has been done by author [8] which may be helpful in doing experiment on multilingual.

From this literature review, it has been observed that lots of challenges have to be taken care of while doing Hinglish Sentiment Analysis like:

- There is no automatic way to convert Hinglish Language sentences to the English language. Many authors have done this work manually.
- Noisy data has to be handled properly before doing Sentiment Analysis.
- Degree modifiers, punctuations, etc. words have to be figured out.
- A very few authors have combined the Vader technique with machine learning techniques. So, one can use this combination for the experiment.
- Phrase, Sarcasm, and irony words to be rendered correctly as they may interpret wrong meaning.

### 3 Proposed Model

For doing Sentiment Analysis, lots of techniques have been available. Numerous researches have been done on different approaches like Machine Learning and Lexicon Based [9]. These techniques do Sentiment Analysis for English Language only.

It has been observed from the researches done by many researchers that Vader Lexicon based techniques give the highest accurate result for finding polarity of English sentences [10, 11]. The reason is that Vader overcomes all the limitations like Noisy data, Punctuations, Capitalization, Degree modifiers, etc. while this limitation is not handled by the most popular SentiWordNet technique [12, 13]. But Vader has been failed to find the correct polarity of sentences other than the English language. In India, many people prefer Hinglish Language to share their feelings. Doing sentiment analysis of these Hinglish sentences (tweets) becomes a challenging task. Still, there is a wide range of different models to overcome this problem. But there is no standard algorithm. So, a new automated model has been proposed using python 3 for doing Sentiment Analysis of Hinglish sentences.

### 3.1 Collection of Raw Data

Data (Tweets) has been scooped from Twitter using Twitter API. The scraper has been written in Python Language for streaming data. Around 8500 tweets (Dataset) have been fetched from Jan 2021 to July 2021 on the topic “Delhi University”. This period is a remarkable time for COVID19. So, we got a large number of unbalanced dataset based on students liking and disliking online classes instead of offline classes.

### 3.2 Description for New Model

An algorithm for new proposed SATV model is as follows:

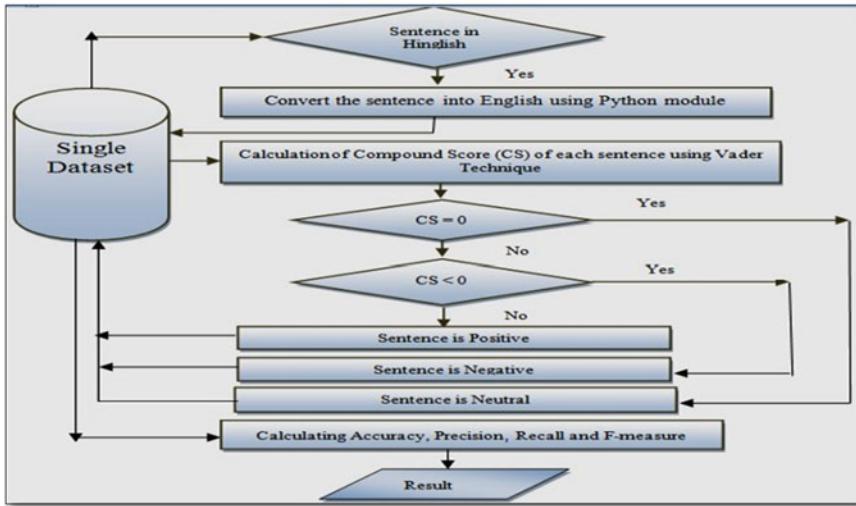
```

Input: Large Dataset containing tweets
Output: Polarity of each tweet and Calculation of accuracy
1 Read each tweet from corpus
2 For each tweet do {
3   IF (Sentence in Hinglish language) THEN {
4     Convert the sentence in English language using Python module }
5   ELSE {
6     No change for English sentence }
7   END IF
8   DO {
9     Calculate positive, negative, neutral and compound scores using Vader
      technique.
10    IF (Compound score < 0) THEN {
11      Polarity = Negative }
12    ELSEIF (Compound score > 0) THEN {
13      Polarity = Positive }
14    ELSE {
15      Polarity = Neutral }
16    END IF
17    Save Polarity to dataset in new field named "polarity"
18  }
19  WHILE end of dataset
20 END
21 Read whole Dataset after finding polarity from above steps.
22 Split Dataset in the ratio of 40:60. 40 for testing and 60 for training.
23 Calculate accuracy using Naïve Bayes and Linear SVC techniques.

```

In the above algorithm, each sentence of the dataset has been checked if the sentence is in Hinglish language then the sentence has been converted into English language using the python 3 translator module. After that, the compound score of each data has been calculated. If the compound score is of positive count then the sentence is treated as positive polarity. Similarly, if the compound score is of negative count then the sentence is treated as negative polarity. Otherwise, the sentence is of neutral polarity. After this step, the data has been split into 40:60 ratios for testing and training respectively. Finally, Naive Bayes and Linear SVC (machine learning techniques) have been used for finding accuracy. The Block diagram for the SATV model is as shown in Fig. 1.

Comparison of SATV model has been done with Vader Technique to find out advantages of SATV model over Vader Technique. Table 1 shows the comparison of the Vader Technique with the SATV model.



**Fig. 1** Block diagram for SATV model

Above Table 1 show that Vader Technique failed to find out the correct polarity of the sentences (Tweets). The reason is that Vader is unable to understand the Hinglish language words and sentences hence always giving Neutral polarity whereas, the SATV model gives correct polarity of all Hinglish sentences. So, the SATV model seems to be more accurate than Vader Technique.

Now accuracy using Naïve Bayes and Linear SVC has been done. Also, confusion matrix has been used for finding Accuracy, Precision, Recall, and F-measure. This result has been compared with the Vader technique to check the performance of the proposed SATV model. Figure 2 shows the comparison between them.

The above figure shows that the SATV model gives high Accuracy, Precision, Recall, and F-measure results than Vader Technique.

## 4 Result

This experiment shows an overview of the performance of the new proposed SATV model. It has been observed that Vader failed to find the correct polarity of the Hinglish sentence (Tweets). But SATV model gives the correct polarity of Hinglish sentences (Tweets). Hence, overall accuracy has been increased by using this new model. Vader gives a good accuracy result that is between 87 and 96%. But still, there is scope for improvement. It has been observed that Vader always gives neutral polarity to Hinglish sentences. Whereas the SATV model overcomes this limitation of the Vader technique and gives very high accuracy results between 94 and 99%. So

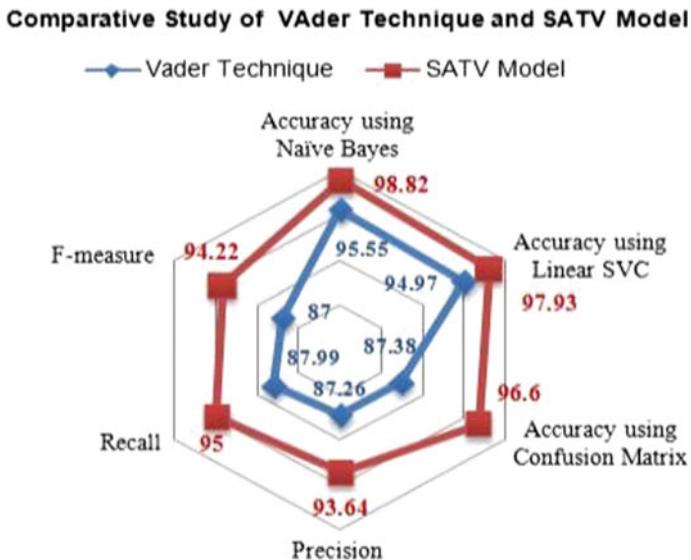
**Table 1** Difference between the Vader Technique and SATV Model on the bases of sentence and their polarity

S. No	Raw sentence (Tweets)	Sentence after using Vader technique	Sentence after using SATV model	Polarity using Vader	Polarity using SATV model
1	Please college khol do.... humme depression ho rhe hai ghar me	Please college khol do.... humme depression ho rhe hai ghar me	Please open college.... We are getting depression in the home	Neutral	Negative
2	Hum Bohut problem face kr rahe hai online classes me	Hum Bohut problem face kr rahe hai online classes me	we are facing a lot of problems in online classes	Neutral	Negative
3	college ka Mera first year campus me aae bina pura ho gaya	college ka Mera first year campus me aae bina pura ho gaya	Completed my first year of college without even visiting the campus	Neutral	Negative
4	shiksha ke baare mein PM Modi ke ghoshana ke anusar kuch classes shuru ke ja sakte hai	shiksha ke baare mein PM Modi ke ghoshana ke anusar kuch classes shuru ke ja sakte hai	As per the announcement by PM Modi about education can be started same classes	Neutral	Positive
5	I want to kill myself lekin maine iss semester ki fees dedi hai:)	I want to kill myself lekin maine iss semester ki fees dedi hai:)	I want to kill myself but I have given fees of this year:)	Neutral	Positive
6	Delhi University science students ke lie phir se khulega	Delhi University science students ke lie phir se khulega	Delhi University reopen for science students	Neutral	Positive

it can be concluded, the SATV model can be recognized for doing Hinglish Sentiment Analysis.

## 5 Conclusion and Future Work

Sentiment Analysis is very popular for finding the sentiment of people in various fields like politics, Education, Products, etc. In this experiment, we have taken 8500 Delhi University tweets. These tweets consist of student opinions for online classes instead of offline classes in the Covid19 pandemic period. It has been observed that



**Fig. 2** Shows the comparative study of Vader technique and SATV model

students are not happy with the online classes and they want colleges to be open as soon as possible.

A new automated model named SATV (Sentiment Analyzer Translating Vader) has been proposed for doing Sentiment Analysis of Hinglish Language. This SATV model has been compared with the Vader technique and it has been observed that SATV model gives high Accuracy, Precision, Recall, and F-measure values than Vader Technique.

This experiment has been done for University data but the same experiment can be done on other datasets like transportation, politics, etc. Also, data can be taken from other social media or blogging sites. For this experiment, Python has been used other platforms can be used like Hadoop, Weka, etc. Also, other techniques like MultinomialNB, Logistic Regression, and, etc. can be used for finding accuracy.

## References

1. M.M. Agüero-Torales, J.I. Abreu Salas, A.G. López-Herrera, Deep learning and multilingual sentiment analysis on social media data: An overview. *Appl. Soft Comput.* **107** (2021)
2. S. Kaur, G. Singh Josan, Sentiment analysis of code-mixed language. *Int. J. Adv. Sci. Technol.* **30**(1):1–11 (2021)
3. D. Gamal, M. Alfonse, E.-S. M.El-Horbaty, A.-B. M.Salem, Twitter benchmark dataset for Arabic sentiment analysis. *Int. J. Mod. Educ. Comput. Sci.* **11**(1):33–38 (2019)
4. K. Singh, Lexicon based sentiment analysis for Hindi reviews. *IJARCCE* **10**(1), 2278–1021 (2021)

5. S. Seshadri, A. Lohidasan, R. Lokhande, A. Save, T. PNagarhalli, A. Professor, A new technique for opinion mining of Hinglish words. *Int. J. Innov. Res. Sci. Eng. Technol.* (2007)
6. P. Singh, E. Lefever, Sentiment analysis for Hinglish code-mixed tweets by means of cross-lingual word embeddings, in *European Language Resources Association (ELRA)* (2020)
7. V. Thakur, R. Sahu, S. Omer, Current state of Hinglish text sentiment analysis. *Int. Conf. Innov. Comput. Commun.* **1** (2020)
8. S. Kumar et al., Review of sentiment analysis: A multilingual approach. *Int. J. Adv. Res. Comput. Commun. Eng.* **9**(1) (2020). <https://doi.org/10.17148/IJARCCE.2020.9110>
9. M.D. Devika, C. Sunitha, A. Ganesh, *Sentiment Analysis: A Comparative Study On Different Approaches*, vol 87 (Elsevier, 2016), pp. 44–49
10. S.S. Hanswal, A. Pareek, A. Sharma, Twitter sentiment analysis using rapid miner tool. *Int. J. Comput. Appl.* **177**, 44–50 (2019)
11. S. Singh, A. Pareek, Improving public transport services using sentiment analysis of twitter data. *J. Inf. Comput. Sci.* **10**, 234–250 (2020)
12. M.B. Myneni, L.V.N. Prasad, J.S. Devi, A framework for sementic level social sentiment analysis model. *JATIT* **96**(16), 1992–8645 (2017)
13. S.S. Hanswal, A. Pareek, G. Vyas, A. Sharma, *Sentiment Analysis on E-Learning Using Machine Learning Classifiers in Python* (Springer), p. 1187. 978-981-15-6014-9\_1 (2021)

# Characterization of Molecular Dynamic Trajectory Using K-means Clustering



Rajneet Kaur Bijral, Jatinder Manhas, and Vinod Sharma

**Abstract** Conformations of kinase obtained from molecular dynamic (MD) simulation plays an important role in molecular docking experiment in the field of drug discovery and development. The scanning of all the MD conformations against millions of drug like molecules is not feasible as it requires high computational cost. Clustering techniques have been widely explored to reduce the conformations into manageable size. Clustering is the Artificial Intelligence based technique in which conformations are partitioned into clusters that exhibit similar behavior thus reducing the scanning cost. This paper analyzed two techniques Principal component analysis and K-means clustering to analyze the molecular simulation data. The clusters thus obtained are evaluated by using Average Silhouette Width and Calinski-Harabasz Index by considering varying number of clusters from 2 to 10 by varying Principal Components from 4 to 6. The result obtained by the proposed framework shows that Average Silhouette Width for K-means clustering is independent of Principal Component subspace.

**Keywords** Molecular dynamics simulation · K-means clustering · Principal component analysis

## 1 Introduction

Epidermal growth factor receptor (EGFR) [15], a receptor tyrosine kinase physiological function is to regulate epithelial tissue development and homeostasis. Deregulation of the EGFR mainly in lungs and breast cells lead to oncogenesis. By inhibiting the signaling pathway of EGFR kinase, tumors can be treated. Kinase undergoes through enormous conformational changes and these conformational changes can be

---

R. K. Bijral (✉) · V. Sharma

Department of Computer Science and IT, University of Jammu, Jammu, J&K, India  
e-mail: [rajbijral@gmail.com](mailto:rajbijral@gmail.com)

J. Manhas

Department of Computer Science and IT, Bhaderwah Campus, University of Jammu, Bhaderwah, J&K, India

studied by Molecular Dynamic (MD) Simulation. MD simulation [16] gives insight about atomic level structure of the bio molecules that helps to understand the molecular functions and intermolecular interactions. Trajectories thus generated by MD contain thousands of conformations. The scanning of these enormous numbers of conformations with millions of drug like molecules to find the inhibitor for kinase by traditional methods is the long, complex and time consuming. To reduce the time and cost required for the scanning the enormous conformations Artificial Intelligence (AI) techniques are applied, one of this is clustering. In clustering, conformations exhibiting structural homogenous behavior are grouped into one cluster and then the representative of each cluster is scanned with drug like molecule assuming that other conformations in same cluster will exhibit similar behavior. There are different clustering techniques; K-means clustering is one of them. In this paper, we applied Principal Component Analysis (PCA) and K-means clustering on MD simulation trajectory of 1M17. The clusters thus obtained are validated using two validation indexes; Calinski-Harabasz (CH) Index and Average Silhouette Width (ASW) by varying number of clusters from 2 to 10.

## 2 Literature Review

Artificial Intelligence techniques have been explored in various fields [18] including: clinical sector [19, 20], social media [21], E-governance [22], drug discovery [17]. Epidermal growth factor receptor (EGFR) either by mutation or over expression led to different types of cancer in human [1, 2]. Finding out the inhibitor for kinase by traditional methods is very costly and complex problem. Many AI based techniques have been explored to find out the ligand for the kinase to reduce the cost and time required. MD simulation [9] captures the dynamic behavior of the kinase. The conformational changes generated and captured by MD are analyses by using different AI techniques. Different clustering techniques including hierarchical clustering [3], kmeans [4], k-medoids [5] etc. are successfully applied on MD simulations. Abramyan et al. [6] studied clustering analysis taking into consideration both molecular orientation and conformations. They evaluated several agglomerative hierarchical clustering algorithms and find out that ward method as the best performing cluster analysis method in their experiment. Phillips et al. [7] developed model based on polymer theory for clearly-defining dynamic and essential properties of MD simulations. Spectral clustering was applied for the analyses of the disordered protein simulations by comparisons and contrasting of the extracted clusters statistical properties. De Paris et al. [8] developed a framework for clustering trajectories based on the structural features of receptor substrate-binding cavity for the optimization of docking experiments. They used k-means clustering and validated the partitions thus obtained after clustering using validity criteria, and validated their results by interaction of 20 ligands and a fully flexible receptor (FFR) model. From the literature being reviewed, it has been observed that different AI techniques have been used for drug discovery [17].

### 3 Materials and Methods

MD Simulation trajectory of EGFR kinase conducted on Desmond tool was received from D. E Shaw research [14]. Length of the simulation is 41 microseconds long and each frame corresponding to 4th nanosecond was taken for experiment, thus there are 10,279 frames in the trajectory. To study the essential structural dynamic of the conformation Root Mean Square Deviation (RMSD) and Root Mean Square fluctuation (RMSF) have been conducted. PCA has been conducted to extract the structural dynamics of the kinase. K-means Clustering has been used to cluster the conformation exhibiting similar behaviour.

#### 3.1 Principal Component Analysis

Different dimensional reduction Techniques including t-Distributed Stochastic Neighbor embedding (t-SNE) [23], Isomap [24], PCA [12] etc. have been explored in the field of drug discovery for analyzing the data generated by MD simulation. In tSNE high-dimensional points are converted into low dimension points by producing their proximity [23]. Isomap [24] is a non-linear dimensionality reduction technique that is based on spectral theory that preserves the geodesic distance between pairs of points of the dataset while preserving the variability as much as possible. The set of new variables formed are uncorrelated and linear function of original variables and are called principal components (PCs). They are represented by the eigenvectors of the covariance matrix where 1st principal component represent maximum variation and so on. In our Experiment PCA of the cartesian coordinated of the EGFR kinase is done to record the essential dynamic in the trajectory.

#### 3.2 K-means Clustering

Clustering is the unsupervised learning algorithm in which data having common characteristics are grouped into a cluster or group. Depending upon measures and initial conditions there are different types of clustering algorithms including: Partitional clustering [26], Hierarchical clustering [3] and density based clustering [25]. In partitional clustering the different clusters are generated based on similarities. This algorithm requires number of the clusters to be generated in advance. Some of the partitional based algorithms are: K-means, K-medoids and CLARA algorithm. In hierarchical clustering data is grouped into tree of clusters. In this clustering technique initially every data point is treated as separate cluster and they are merged based on some similarity measures until they form single cluster. Density based clustering distinct clusters are identified on the bases of data point density. In this study the Kmeans clustering has been used.

K-means [13] clustering is the unsupervised learning algorithm that partitioned data into distinct clusters such that data belongs to only one group based on some similarity measures. In this algorithm number of clusters is the predetermined and data is divided to minimize the distance between the data within the same clusters.

### 3.3 Clustering Validation Criteria

**Average Silhouette Width:** Average Silhouette Width (ASW) [10] is used to measure the clustering quality. It computes the average silhouette for the observation for different number of clusters. The result that maximizes the average silhouette value gives us the optimal number of clusters.

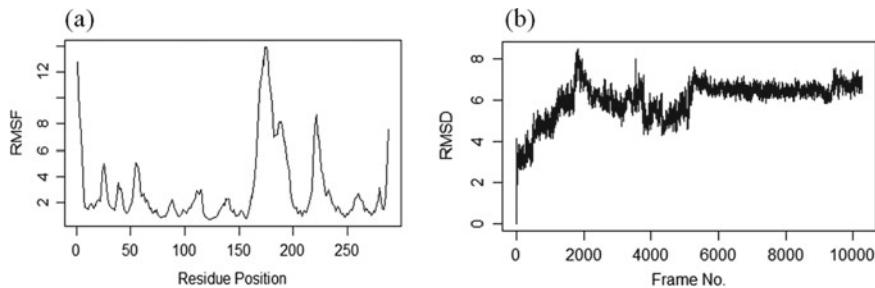
**Calinski-Harabasz (CH) Index:** Calinski-Harabasz (CH) Index [11] is the validation index to evaluate the compactness within the cluster and separation between the clusters. The index is defined as the ratio of variation between clusters and within cluster variation. Calinski-Harabasz Index is defined as:

$$CH(K) = [B(K)/W(K)]X[(n - K)/(K - 1)]$$

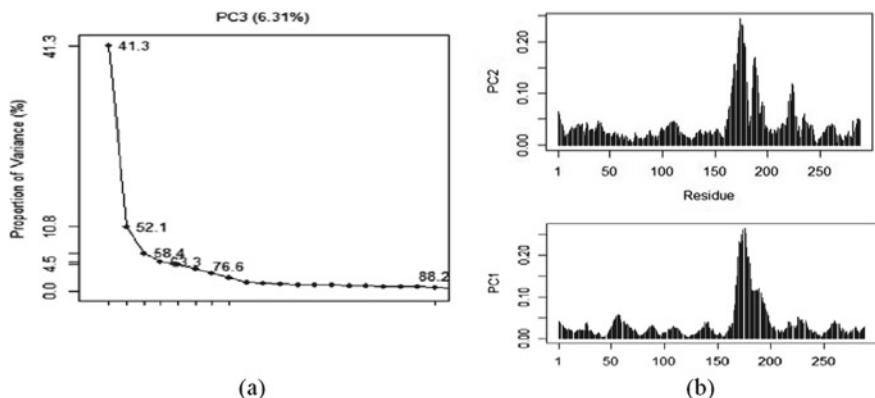
where n is the data point, K is the number of clusters, W(K) is variation within cluster and B(K) is the inter cluster variation. The K value having the highest values gives the optimal number of clusters.

## 4 Result and Discussion

MD simulation of EGFR kinase having 10,279 conformations is saved as 3D frames in trajectory. The overall structure deviation of EGFR over the time is analyzed by RMSF. Considering alpha-carbon, RMSD is done to capture the dynamics of the kinase. RMSF and RMSD of the kinase are shown in Fig. 1. It has been observed from the experiment that maximum RMSD of 8.504° Angstroms and mean deviation of around 6.019° Angstroms. PCA has been applied on the 10,279 frames of trajectory of kinase. PCA analysis plot is shown in the Fig. 2. 1st, 2nd, 3rd, 4th, 5th and 6th PC contribute variance 41.327, 10.779, 6.312, 4.899, 4.462, 3.710 respectively, total variance of 71.488 is captured by first 6 PCs and remaining PCs capture less than 3 percent of variance. The Fig. 2 shows the residue wise contribution in 1st and 2nd PCs. K-means clustering has been applied by considering 4 PCs, 5 PCs and 6 PCs to know affect of PC subspace on k-means clustering. Quality of the clusters is validated by ASW and CH Index by varying clusters number from 2 to 10. The number of optimal clusters obtained by ASW remains same i.e. 4 by varying the PC subspace from 4 to 6. CH Index gives the different result by varying the PC subspace. The Fig. 3 shows the result obtained by ASW and CH Index. It has been observed



**Fig. 1** **a** Root mean square fluctuation **b** Root mean square deviation standard to measure the structural distance between coordinate sets

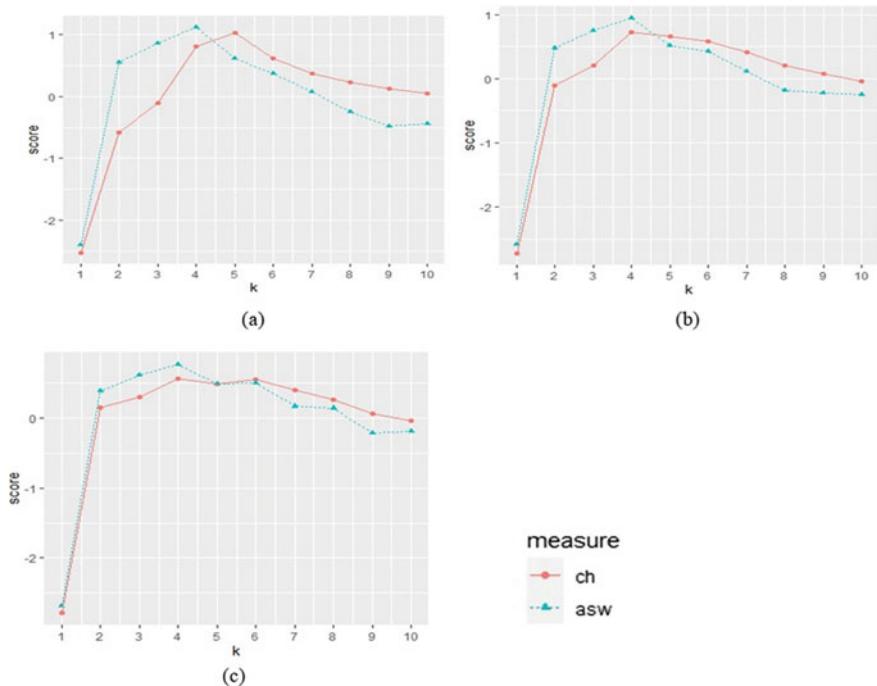


**Fig. 2** **a** Principal component analysis plot on the MD coordinate data of the EGFR (1m17.pdb) and **b** Residue wise contribution of PC1 and PC2

that CH Index is dependent on PC subspace and ASW for K-means clustering is independent of PC subspace.

## 5 Conclusion and Future scope

The framework is developed by applying PCA and K-means clustering on 10,279 frames on EGFR kinase. The clusters thus obtained are evaluated using ASW and CH Index by varying PC subspace from 4 to 6 and considering the number of clusters from 2 to 10. It has been observed that optimal number of clusters thus obtained by ASW remains same thus independent of PC subspace. The optimal number of clusters obtained by using CH Index depends on the number of PC subspace considered. By clustering the conformations into clusters and considering the representative of each cluster for screening against drug-like molecules thus, reduces the computational



**Fig. 3** Evaluation of the partition obtained using K-means clustering algorithm by Average Silhouette Width (ASW) and Calinski-Harabasz (CH) Index. **a** By considering 4 PC subspace **b** By considering 5PC subspace and **c** By considering 6 PC subspace

cost. In future different dimensional techniques and other clustering techniques can be applied on MD simulation data.

## References

- Y. Yarden, G. Pines, The ERBB network: at last, cancer therapy meets systems biology. *Nat. Rev. Cancer*, 553–563 (2012)
- J. Schlessinger, Receptor tyrosine kinases: legacy of the first two decades. *Cold Spring Harb. Perspect. Biol.* (2014)
- F. Murtagh, P. Contreras, *Methods of Hierarchical Clustering. Computing Research Repository—CORR* (2011)
- R. Shang et al., Analysis of simple K-mean and parallel K-mean clustering for software products and organizational performance using education sector dataset. *Sci. Program.* (2021)
- M. Adepeju, S. Langton, J. Bannister, Anchored k-medoids: a novel adaptation of k-medoids further refined to measure long-term instability in the exposure to crime. *J. Comput. Soc. Sci.* 655–680 (2021)
- T.M. Abramyan, J.A. Snyder, A.A. Thyparambil, S.J. Stuart, R.A. Latour, Cluster analysis of molecular simulation trajectories for systems where both conformation and orientation of the sampled states are important. *J. Comput. Chem.* 19,731,982 (2016)

7. J.L. Phillips, M.E. Colvin, S. Newsam, Validating clustering of molecular dynamics simulations using polymer models. *BMC Bioinf.* (2011)
8. R. De Paris, C.V. Quevedo, D.D. Ruiz, O. Norberto de Souza, R.C. Barros, Clustering molecular dynamics trajectories for optimizing docking experiments. *Comput. Intell. Neurosci* (2015)
9. I. Srivastava, A. Kotia, S.K. Ghosh, M.K.A. Ali, Recent advances of molecular dynamics simulations in nanotribology. *J. Mol. Liq.* (2021)
10. F. Batool, C. Hennig, Clustering with the average silhouette width. *Comput. Stat. Data Anal.* (2021)
11. W. Wang, T. Shi, Y. Zhang, Q. Zhu, An affinity propagation-based clustering method for the temporal dynamics management of high-speed railway passenger demand. *J. Adv. Transp.* (2021)
12. M.R. Mahmoudi, M.H. Heydari, S.N. Qasem, A. Mosavi, S.S. Band, Principal component analysis to study the relations between the spread rates of COVID19 in high risks countries. *Alexandria Eng. J.* 457–464 (2021)
13. C. Wu et al., k-means clustering algorithm and its simulation based on distributed computing platform. *Complexity* (2021)
14. Y. Shan, A. Arkhipov, E.T. Kim, A.C. Pan, D.E. Shaw, Transitions to catalytically inactive conformations in EGFR kinase. *Proc. Natl. Acad. Sci. USA* 72707275 (2013)
15. G. Bethune, D. Bethune, N. Ridgway, Z. Xu, Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update. *J. Thorac. Dis.* 48–51 (2010)
16. H. Ode, M. Nakashima, S. Kitamura, W. Sugiura, H. Sato, Molecular dynamics simulation in virus research. *Front. Microbiol.* (2012)
17. R.K. Bijral, I. Singh, J. Manhas, V. Sharma, Exploring artificial intelligence in drug discovery: a comprehensive review. *Arch. Comput. Methods Eng.* (2021)
18. D. Manhas, *Analysis on Design Issues of E-Government Websites of India* (2014)
19. H. Kour, J. Manhas, V. Sharma, in *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 1041–1046
20. R. Gupta, M. Kaur, D. Manhas, Tissue level based deep learning framework for early detection of dysplasia in oral squamous epithelium. *J. Multimedia Inf. Syst.* 81–86 (2019)
21. A.M. Koli, M. Ahmed, J. Manhas, An empirical study on potential and risks of twitter data for predicting election outcomes, in *Emerging Trends in Expert Applications and Security. Advances in Intelligent Systems and Computing*, ed. by V. Rathore, M. Worring, D. Mishra, A. Joshi, S. Maheshwari, vol. 841 (Springer, Singapore) (2019)
22. D. Manhas, V. Mansotra, *Critical Evaluation of e-Government Websites Design* (2011)
23. V. Spiwok, P. Kríž, Time-lagged t-distributed stochastic neighbor embedding (t-SNE) of molecular simulation trajectories. *Front. Mol. Biosci.* 7 (2020)
24. H. Stamatil, C. Clementi, L.E. Kavraki, Application of nonlinear dimensionality reduction to characterize the conformational landscape of small peptides. *Proteins* **78**, 223–235 (2010)
25. J. Sander, in *Encyclopedia of Machine Learning*, ed. by C. Sammut, G.I. Webb, pp. 270–273 (Springer US, 2010)
26. V. Thambusamy, *Applications of Partition based Clustering Algorithms: A Survey* (2013)

# Data Warehouse Modernization Using Document-Oriented ETL Framework for Real Time Analytics



Monika Patel and Dhiren B. Patel

**Abstract** For past several years Relational Databases performing and providing services to many applications. As data is growing, the necessity for a new kind of database which can handle such a huge amount of data in semistructured and unstructured form has also increased. NoSQL databases are able to manage complexity of data structure as well as they can handle such a huge amount of data. Analysis is again a challenging process of such semi-structured and unstructured data. Traditional data warehouses are incapable of analyzing such schema-less data for decision making as relational databases need to know schema in advance. ETL (Extract-Transform-Load) is the process used to collect data from source, process and transform them and storing into a data warehouse for further analysis. The paper presents a framework for an ETL process for document-oriented data warehouse which provides real-time analytics using classification approach at the data warehouse stage. The proposed framework is designed and verified to enhance execution time for real-time analytics.

**Keywords** Extract-transform-load framework · Data warehouse · MongoDB · Real-time analytics · NoSQL/SQL

## 1 Introduction

For analysis and storing data, data warehouses are used from decades in the industries. In the scenario of Big Data DWs become more necessitate to handle semistructured and unstructured data. But traditional Data Warehouses have limitations related to structure of data. NoSQL databases can overcome the limitations of traditional databases and allows schema-free data storage and also provides horizontal scalability. As the data increases substantially, it is important to integrate the data to

---

M. Patel (✉) · D. B. Patel

Department of Computer Science, Gujarat Vidyapith, Ahmedabad, Gujarat, India  
e-mail: [monapatel18r@gmail.com](mailto:monapatel18r@gmail.com)

D. B. Patel

e-mail: [dhiren\\_b\\_patel@gujaratvidyapith.org](mailto:dhiren_b_patel@gujaratvidyapith.org)

extract the vision for business expansion and policy formation for most businesses [1, 2]. Extract-Transform-Load process is used to handle structured, semi-structured and unstructured data for Data Warehouse [3].

### ***1.1 Relational Data Warehouse***

Businesses require to store their data into data warehouses for analytics to take crucial decisions. Typically, data warehouses are designed and managed using relational databases for knowledge discovery using relational queries. The extensive growth of unstructured data needs expansion of a data warehouses that can manage data in an unstructured way, with each record having different attributes. Non-Relational technologies have ability to accumulate and process such schema free data and provides scope to design new age data warehouse [4].

### ***1.2 ETL***

For creation of a Data Warehouse, ETL (Extract-Transform-Load) is the primary process for collecting and organizing data for business decisions. Extracting, Transforming and Loading is cumbersome process in the designing of any data warehouse. An Extraction is the process of collecting data from sources and temporary lodging into the stage area for further transformation. Transformation includes cleaning, splitting, joining, and aggregating data. Lodging process loads data into the designated Data Warehouse [5, 6].

### ***1.3 NoSQL Based Data Warehousing***

NoSQL databases are new age databases and has ability to manage huge amount of data in unstructured form. Researchers have identified and pinpointed many advantages of NoSQL databases over relational databases and emphasized the scope of data warehouse development using non-relational databases. Some of the foremost benefit of NoSQL databases are schema-free data structure and scalability [7].

### ***1.4 Real-Time Analytics***

Real-time analytics is the need for current businesses to make quick decisions to cop-up the competition. Many researchers are working with the real-time analytics using NoSQL technologies [8–10]. Papers deals with real-time stream processing using

stream join algorithms to improve latency. In another paper, authors have designed a framework for handling structured and unstructured data using NoSQL technologies for data analytics.

The proposed paper provides the modern framework of handling structured and unstructured data using modern technologies (NoSQL) for data analytics [4, 11]. Prototyping of the proposed framework with execution time and memory usage will be analysed to check the performance.

The paper organizes into sections, where Sect. 2 represents related work done in the proposed research area. Section 3 shows the Proposed Framework, Sect. 4 represents methodologies and process will be used to execute the proposed framework. Section 5 covers Results and Discussion of the proposed framework and Sect. 6 Concludes the paper.

## 2 Related Work

The authors have analyzed 15 most prevalent ETL tools and carried out main features an ETL tool can provide. The paper also suggests the appropriate ETL tool for a specific type of organization. Further, it has highlights the limitations of most ETL tools such as horizontal scalability, real-time processing and automation [1]. The researchers analyzed the performance of incremental lodging in ETL using codebased ETL tools. Incremental loading is done by snapshot-based Change Data Capture, which will record the changes in the data while loading them into the data warehouse, then dimension and fact processing will be performed. Furthermore, authors have developed a code-based ETL framework to supports RDBMS and NoSQL database with rich transformation techniques [12].

In this paper, a novel data warehouses approach is presented using document-oriented databases. The approach consists of five steps such as choosing NoSQL database, extracting schema, identifying structure graph, multidimensional schema and designing the data warehouse schema [13]. In this research, authors have provided a data warehouse solution for NoSQL and SQL data sources for improved stream processing using MongoDB. The analysis between structured and unstructured semi-stream join processing using MongoDB has been made while implementing Extract-Transform-load phase. CPU and memory usage is analyzed while executing using synthetic and real datasets. Finally, the authors have guided for choosing appropriate memory size to reduce disk I/O cost for effective implementation of real-time data warehouse [14]. This paper includes has reviewed 74 articles for real-time stream processing systems in systematic way, which provide guidance to researchers for implementation of a robust real-time stream processing framework for real-time Data Warehouse. This study pinpoints a problem of less research done for ETL for unstructured data [8].

The authors have designed an ETL based platform for transformation of multi-dimensional concept-based model using document-oriented database for NoSQL

warehousing. The Business Process Modeling Notation is used for modeling transformation rules and then the model is evaluated for read and write request latency using TPC-DS benchmark. Based on the results researchers have come to the conclusion that NoSQL data warehouse manage data volume and queries without affecting performance [15]. This research proposes BigDimETL with multidimensional structure for the ETL process to achieve knowledge discovery. To develop a data warehouse MapReduce and Hbase are used to store data into facts and dimensions. Three queries are used to analyze the execution time, queries can be run using different factors such as table size, joins, aggregations and concurrent queries [16]. The researchers have highlighted the need of tools and techniques for unstructured data, as available tools are able to manage structured data only. The paper designed the method for handling unstructured data using NoSQL platform like MongoDB. Researchers have focused on features of MongoDB, which can analyze data of any structure without using classy data warehouse workloads [17].

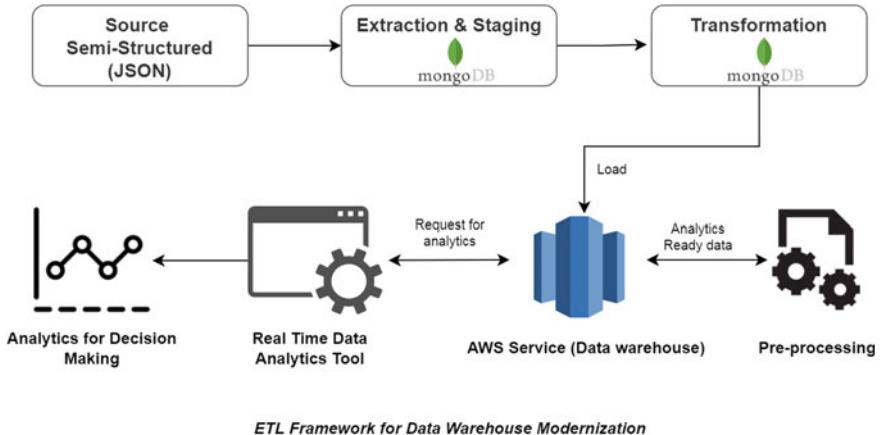
### 3 Proposed Work

Prevailing ETL frameworks are usually transform data into structure form before storing them into Data Warehouse for analysis. The proposed framework has been designed based on literature review done from similar papers of ETL frameworks [1]. Transforming semi-structured and unstructured data into structured data to create a data warehouse is inefficient and needs the schema to be known in beforehand.

The present Semi-structured and unstructured data gathered from various applications makes it hard to transform data into the structure form for analysis [18].

The proposed ETL framework is designed using a Non-Relational Document-Oriented Database for modern Data Warehouse. Figure 1 represents proposed ETL framework. Very less work done in the field of ETL process to store and analyze documents (JSON) in data warehouse.

Every business in the market is facing blast of data because of wide use of internet and online platforms. It is very important to store and analyze products, sales and customer data to understand the pattern and to increase growth of businesses. The data generated from online platforms have no structure or have document like structure with no schema defined. NoSQL databases are becoming popular to manage such schema-free data with and also provides horizontal scalability. To encounter the organizational requests of handling semi-structured data for analysis the proposed ETL framework is designed. The proposed framework challenges typical ETL processes for storage, retrieval, scalability and real-time analytics [19, 20]. The proposed framework receives input from source (JSON documents) into MongoDB for transformation. MongoDB's aggregation framework creates pipeline for data aggregation. Transformed documents will be stored using Amazon Web Services as a Data Warehouse [21]. To enhance real-time analytics, data pre-processing will be performed on transformed data for on-demand analytics using data classification approach.



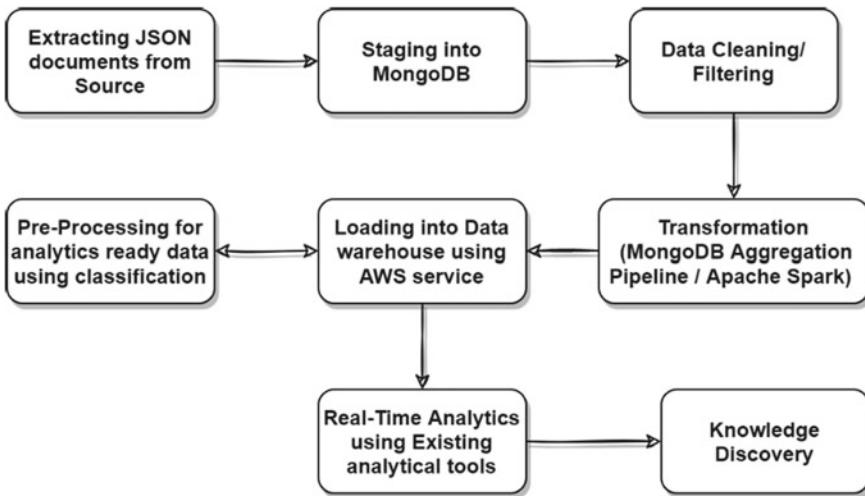
**Fig. 1** Proposed ETL Framework

Selected classification method will pre-process the data into analytics-ready data for quicker results [11]. Apache Spark data stream processing engine can be used to speed up the process and to identify the essential components while analytics [22, 23]. Pre-processed data will reduce the analysis time and provides real-time analytics. An ETL processing tool HCI integrates health care databases into knowledgebase using AWS cloud services [24].

## 4 Methodology

Figure 2 shows the method of implementation stated in ETL framework.

- Step 1** Extraction of JSON documents from source system. GSTR1 and GSTR 2A files of various clients are extracted from Goods and Service Tax to analyze the sales and purchase over the time.
- Step 2** Staging into MongoDB for further process. In MongoDB JSON documents will be stored as collections.
- Step 3** Data Cleaning, Filtering, attribute selection. This step selects the required attributes for further analysis.
- Step 4** Transformation of documents using MongoDB aggregation Framework. Using Aggregation pipeline data are transformed and grouped using “\$unwind”, “\$group” command for data aggregation.
- Step 5** Loading into Data Warehouse using Amazon Web Services.
- Step 6** To Speed up the analysis, Pre-processing is performed using classification approach to classify the data into categories and understand the relationship between variables.
- Step 7** Processed data will be stored into Data Warehouse to fasten the analysis.



**Fig. 2** Flow of proposed ETL Framework

**Step 8** Testing Execution time using Existing Real-Time analytics tools.

**Step 9** Knowledge Discovery and decision-making using analytics.

## 5 Results and Discussion

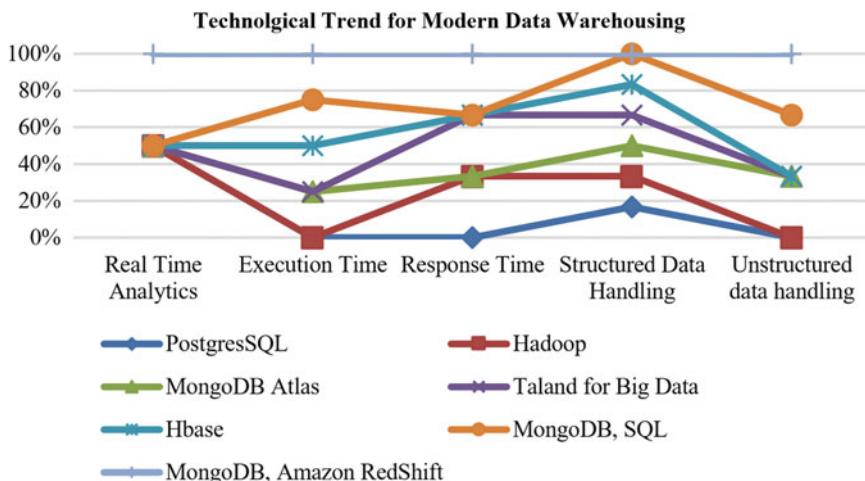
As can be seen, the approaches and technologies used to manage unstructured data is limited and they only provide solution till one or more extent of the ETL and Data Warehouse process (Table 1).

The studied approaches can provide transformation of unstructured data into structured and vice versa to design a NoSQL based Data Warehouse for Real-Time Analytics. Analyzing unstructured data from the Data Warehouse can be more complicated, thus it requires analytical techniques to speed up the process. The proposed work covers major aspects of designing, implementing and providing real-time analysis using classification approach at the data warehouse for unstructured data.

Figure 3 shows the technologies used for designing NoSQL based data warehousing using structured/unstructured data and points out the performance at the specific phase of implementation. Which leads to further research and development in the area using other technologies. Here majorly used technologies are evaluated based on capacity and performance for structured data handling, unstructured data handling, real-time analytics, execution time and response time. Proposed work overcomes the limitations in the work done till the date with new trend of technologies with added data mining approach at the data warehouse level for real time analytics.

**Table 1** Technological trend in the development of NoSQL based Data Warehouse

References number	Input data	Technology	Scope of the research
[12]	Database tables, CSV files	Postgres SQL Anaconda Python	Modeling of a near real time process
[13]	DBLP, LINKEDIN	Hadoop MapReduce	Design a data warehouse Scheme from unstructured data
[14]	100YWeathe rSmall	MongoDB Atlas	Semi-stream join processing at ETL phase
[15]	Multi-dimensional schema	Talend for big data, Java routines	Transformation rules to transform multi-dimensional schema into document-oriented system
[16]	Twitter	Hbase, Oracle SQL, Hadoop hortonworks	Adapting ETL process with big data for decision making
[17]	CSV, MP3 file	MongoDB, SQL	Query processing comparison in SQL and MongoDB
Proposed work	GST—JSON files	MongoDB, Amazon red shift	ETL framework for document oriented data warehouse with real-time analytics using classification approach

**Fig. 3** Technological Trend for Modern Data Warehousing

Appling Data mining on the data warehouse compliments each other by improving real-time analytics.

## 6 Conclusion

This paper proposed an ETL framework for data warehousing using NoSQL databases. The framework will improve execution time by pre-processing data priorly for better performance in the era of big data. This newly designed framework using NoSQL technologies seems to offer promising results for handling huge amount of semi-structure data and real-time analytics. The study of latest technologies has been done to identify the superlative tools which can improve performance in every aspect of developing and executing document-oriented data warehousing. The researcher has analyzed and determined that the working on semi-structured data directly will save time of transformation into structured form. As future work, we will implement the proposed ETL framework with defined methodologies which will establish transformation at the data warehouse stage using appropriate classification approach before analytics.

## References

1. M. Patel, D.B. Patel, *Progressive Growth of ETL Tools: A Literature Review of Past to Equip Future*, vol 1187 (Springer Singapore, 2021)
2. F. Halawa, H. Dauod, I.G. Lee, Y. Li, S.W. Yoon, S.H. Chung, Introduction of a real time location system to enhance the warehouse safety and operational efficiency. *Int. J. Prod. Econ.* **224** (2020)
3. M.Y. Santos, C. Costa, C. Andrade, *Enhancing Big Data Warehousing for Efficient, Integrated and Advanced Analytics*: Visionary Paper The final authenticated version is available online at [https://doi.org/10.1007/978-3-030-21297-1\\_19](https://doi.org/10.1007/978-3-030-21297-1_19) (June, 2019)
4. A. Sabtu et al., The challenges of extract, transform and loading (ETL) system implementation for near real-time environment. *Int. Conf. Res. Innov. Inf. Syst. ICRIIS*, 3–7 (2017)
5. S. Gorhe, *ETL in Near-Real Time Environment: Challenges and Opportunities*
6. R. Bagave, *Enhancing Extraction in ETL flow by modifying as P-ETL based on Spark Model Msc in Cloud Computing Rahul Bagave Supervisor*
7. S. Muhammad, F. Ali, *Next-generation ETL Framework to address the challenges posed by Big Data*
8. E. Mehmoor, T. Anees, Challenges and solutions for processing real-time big data stream: a systematic literature review. *IEEE Access* **8**, 119,123–119,143 (2020)
9. X. Li, Y. Mao, Real-time data ETL framework for big real-time data analysis, in *2015 IEEE International Conference on Information and Automation ICIA 2015—Conjunction with 2015 IEEE International Conference on Automation and Logistics*, Aug (2015), pp. 1289–1294
10. R.J. Santos, J. Bernardino, M. Vieira, 24/7 real-time data warehousing: A tool for continuous actionable knowledge, in *2011 IEEE 35th Annual Computer Software and Applications Conference* (July) (2011), pp. 279–288
11. E. Zdravevski, P. Lameski, A. DImitrevski, M. Grzegorowski, C. Apanowicz, Cluster-size optimization within a cloud-based ETL framework for Big Data, in *2019 IEEE International Conference on Big Data (Big Data)* (2019), pp. 3754–3763

12. N. Biswas, A. Sarkar, K.C. Mondal, Efficient incremental loading in ETL processing for real-time data integration. *Innov. Syst. Softw. Eng.* **16**(1), 53–61 (2020)
13. S. Bouaziz, A. Nabli, F. Gargouri, Design a data warehouse schema from document-oriented database. *Procedia Comput. Sci.* **159**, 221–230 (2019)
14. E. Mahmood, T. Anees, Performance analysis of not only SQL semi-stream join using MongoDB for real-time data warehousing. *IEEE Access* **7**, 134215–134225 (2019)
15. R. Yangui, A. Nabli, F. Gargouri, *ETL Based Framework for NoSQL Warehousing*, pp. 0–14, Aug (2017)
16. H. Mallek, F. Ghazzi, O. Teste, F. Gargouri, BigDimETL with NoSQL Database. *Procedia Comput. Sci.* **126**, 798–807 (2018)
17. R. Aluvalu, M.A. Jabbar, Handling data analytics on unstructured data using MongoDB. *IET Conf. Publ.* **2018** (CP747) (2018)
18. D. Prakash, *NOSOLAP: Moving from Data Warehouse Requirements to NoSQL Databases*, no. Enase (2019), pp. 452–458
19. R. Indrakumari, T. Poongodi, P. Suresh, B. Balamurugan, The growing role of integrated and insightful big and real-time data analytics platforms, 1st ed., vol. 117, no. 1. Elsevier Inc. (2020)
20. A. Wibowo, Problems and available solutions on the stage of extract, transform, and loading in near real-time data warehousing (a literature study), in *2015 International Seminar on Intelligent Technology and Its Applications (ISITIA) 2015* (2015), pp. 345–349
21. N. Fikri, M. Rida, N. Abghour, K. Moussaid, A. El Omri, An adaptive and realtime based architecture for financial data integration. *J. Big Data* **6** (1) (2019)
22. F.K. Putri, J. Kwon, A distributed system for finding high profit areas over big taxi trip data with MognoDB and Spark, in *2017 IEEE 6th International Congress on Big Data (BigData Congress) 2017* (2017), pp. 533–536
23. A. Priya, K.R. #1, S. Rani, *A Survey Paper on the Comparison of NOSQL Engines (Mongo dB vs. Cassandra) using Spark*, vol. 6, no. 5 (2019), pp. 233–236
24. S.A. Trivedi, M. Patel, S. Patel, Health care cube integrator for health care databases. *Web Semant. Acad. Press.* (2020), pp. 129–151

# Opinion Mining for Breast Cancer Disease Using Apriori and K-Modes Clustering Algorithm



S. Balakrishnan<sup>ID</sup>, K. Suresh Kumar<sup>ID</sup>, M. Balasubramanian,  
and Y. Rajkumar<sup>ID</sup>

**Abstract** Data mining procedures have been broadly used to mine learned data from medicinal information bases. Sentiment Mining is a procedure of programmed extraction of learning by method for the conclusion of others about some specific item, theme or Issue. Sentiment analysis implies deciding the subjectivity, extremity (positive/negative) and extremity quality and so forth, with a bit of text. Clustering is the methodology of making a get together of dynamic things into classes of near articles. In this paper, we are proposing a two way clustering algorithm for breast cancer disease. Apriori hybrid algorithm and K-Modes Algorithm is used to cluster the opinions effectively and to improve the performance in the online data set. Apriori-Hybrid is the mix of count Apriori and Apriori-TID, which can mastermind the huge item sets and improve the accuracy of collection of dangerous development. It can moreover uncover understanding into the basic part that enables each malady type to suffer and thrive, which in this way help in the early revelation of the threatening development. We propose Apriori-Hybrid as an extemporized calculation for tumour characterization.

**Keywords** Clustering · Opinion mining · Breast cancer · K-modes · Apriori hybrid

---

S. Balakrishnan (✉)

Sri Krishna College of Engineering and Technology, Coimbatore, India  
e-mail: [balkiparu@gmail.com](mailto:balkiparu@gmail.com)

K. Suresh Kumar

Saveetha Engineering College, Chennai, India

M. Balasubramanian

Vivekananda Institute of Professional Studies, New Delhi, India

Y. Rajkumar

Bannari Amman Institute of Technology, Sathyamangalam, India

## 1 Introduction

The tumour everything considered grows once cells of a part of the body start to escape directing. These extra cells make a mass of tissue, recommended as new development or neoplasm. Tumours are amicable or dangerous. The science whose objective is to perform depiction of things into a course of action of classes or portrayals is proposed as Pattern Recognition. Things might be “pictures, signal waveforms or any assortment of measures that should be mentioned” [1].

The monstrosity of the assessment is that; “the carcinoma infers remarkable malignancies that make in one or every chest, and it is the most notable sort of contamination among women in made nations”. As indicated by the U.S. Damage Society, one out of eight women can make carcinoma all through their period. The Issue with carcinoma obvious proof is that paying little mind to radiographic chest imaging and screening has thought about a ton of right I.D. of carcinoma, 100% to a half-hour of sabotaging cases don’t emit an impression of being perceived for various reasons. There are fundamentally two sorts of messes up typical in looking at mammograms.

They are “False Positives (F.P.) and False Negatives (F.N.)”. The Computer-Aided undertaking (CAD) will lessen each F.P. besides the F.N. undeniable proof rates. P.C. supported plan is the use of model affirmation going for helping experts in settling on indicative decisions. Recognizing confirmation tasks is moulded by the authority. Our point is to utilize a model affirmation structure to assist restorative experts with a “Second” supposition by closing the mammographic mass options that the lion’s offer exhibits danger [2].

Development is generally a hazardous cell that transforms into a colossal clarification for the destruction and is barely turned away [3, 4]. India is creating carcinoma scourges with a growing extent of more young ladies changing into helpless to the prosperity. In remedial spaces, data getting ready systems are extending “quickly because of the advancement feasibility of those approaches to managing request and conjecture structures, new and novel assessment orientation are known for more clinical and test investigation”. By and large, infection assessment relies on real models that fail to reach to spread in therapeutic as an outcome of the usage of these devices isn’t fits in with a remedial gathering. A learning course of action strategy using information obtained from comprehended data has been among the first genuinely analyzed subjects in quite a while. There are a couple of techniques to predict and arrange carcinoma plans. Most women have more than one perceived threat issue for carcinoma, regardless, can ne’er get the prosperity. The primary ordinary peril components “for carcinoma isn’t solely being polite” and getting more settled. There is moreover more than one clarification for carcinoma.

These could be:

- “Being a woman
- Getting more seasoned
- Having a genetic transformation inside of the brca1 or brca a couple of carcinoma cistron
- Lobular disease in situ (LCIS)

- An individual history of bosom or female inside conceptional organ disease
- A case history of bosom, female inside a conceptional organ or glandular disease
- Having high bosom thickness on a mammogram
- Having a past indicative examination demonstrating atypical dysplasia
- Starting change of life when age fifty-five
- Never having adolescents
- Having your first youngster when age thirty-five
- Radiation presentation, continuous x-beams in youth
- High bone thickness
- Being overweight when the change of life or putting on weight as a grown-up". Clustering is a standout amongst the most helpful errands in the information-digging process for finding aggregates and distinguishing fascinating dispersions and examples in the hidden information. The clustering issue is about parcelling information set into gatherings (groups) such that the information focuses in a group are more like each other than focuses in various bunches [5].

The K-means clustering calculation can't bunch all-out information on account of the uniqueness measure it employs. The "K-modes clustering calculation depends on Kmeans worldview yet evacuates the numeric information constraint whilst safeguarding its effectiveness" [6]. The K-modes calculation stretches out K-means worldview to bunch all-out information by evacuating the impediment forced by K-means through after changes:

- "Using a straightforward coordinating divergence measure or the hamming separation for all-out information objects
- Replacing method for bunches by their modes".

## 2 Review of Literature

Huang [7] introduced "two techniques for instatement for unmitigated information for K-mode grouping calculation and demonstrated that if differing beginning modes are picked, it could prompt better bunching results". Sun et al. [8] proposed an iterative technique because starting focuses refinements calculation for downright information grouping to the setting up of the underlying guides to guide the straight out information sets to bunching results that have better consistency rates. Bradley and Fayyad's "iterative beginning point refinement calculation [9] to the K-modes grouping to enhance the precision and reiteration of bunching results". They utilized a subinspecting strategy to convey the bunching a few times so that the impact of skewed disseminated information ought not to influence the last grouping results. Khan et al. [10] displayed a calculation to figure introductory modes using "Density based Multiscale Data Condensation". Kant et al. [11] displayed an "Automatic and Stable Clustering Algorithm for numerical grouping information". All the more, as of late, Fred et al. [12] utilized the thought of proof bunching to consolidate the consequences of different clusterings (N times) into a solitary information allotment by

surveying every grouping result as a free method confirmation of information association. They did as such by running a K-implies calculation commonly with various parameters or introductions. First, the information is divided into countless and little groups; irregular statements of the K-implies calculation get distinctive deteriorations. The last information segment is gotten by bunching this new comparability lattice, relating to the converging of the group [13].

Topchy et al. [14] introduced a calculation to join different feeble bunching and figured that consolidated grouping gets identical to bunching a straight out information because of some picked accord capacity. Sudha [15] offers “an inspiration regarding major critical diseases and their identification using data processing with a minimum range of attributes and creates awareness regarding diseases that end in death”. Chaurasiaet et al. [16, 17] conducted “a study on the prediction of coronary failure risk levels of the heart disease information with data processing technique like Naïve Bayes, J48 call tree and bagging approaches and CART, ID3 and decision table”.

### 3 Breast Cancer Dataset

The information used as a part of this study is given by the “University Medical Center, Institute of Oncology, Ljubljana, Yugoslavia. The dataset has ten characteristics and aggregate 286 columns,

1. Age: patient’s age at the season of finding;
2. Menopause: menopause status of the patient at the season of finding;
3. Tumor size: tumour size (in mm);
4. Inv-nodes: range 0–39 of axillary lymph hubs indicating bosom tumour at the season of histological examination;
5. Node Caps: infiltration of the tumour in the lymph hub container or not;
6. Level of harm: extent 1–3 the histological evaluation of the tumour. That is evaluation: (1) prevalently that comprise of growth cells, grade: (2) neoplastic that comprises of the standard qualities of malignancy cells, grade: (3) predominately that comprise of cells that are very influenced;
7. Bosom: bosom tumour might happen in either bosom;
8. Bosom quadrant: if the areola considers as an essential issue, the bosom might be isolated into four quadrants;
9. Illumination: patient’s radiation (x-beams) treatment history;
10. Class: no-repeat or repeat depending returning side effects of bosom malignancy in the patients after treatment”.

## 4 Proposed Work

Our proposed approach combines the Apriori hybrid algorithm and k-modes algorithm to improve the clustering accuracy of the breast cancer data set. A large portion of the clustering calculations embraces the thorough inquiry strategy exhibited in the popular APRIORI calculation to find the guidelines and require various information bases [18–26].

The Apriori figuring tends to the going with indispensable Issues [27]. The Apriori produces the sure itemsets by joining the expansive itemsets of the past pass and destroying those subsets which are little in the past desert considering the exchanges the information base. By basically considering colossal itemsets of the past pass, the measure of competitor huge itemsets is reduced on an extremely fundamental level. The found broad itemsets of the key pass are utilized to make the sure blueprints of the resulting pass utilizing() as far as possible. In the major pass, the itemsets with only things are checked. In the third pass, the gigantic itemsets of the resulting pass are considered the applicant sets to find wide itemsets of this pass. When the bright itemsets are found, their sponsorships are numbered to find the liberal itemsets of size two by filtering the information base. This iterative method closes when no new liberal itemsets are found.

### 4.1 *Apriori-Tid Algorithm*

Apriori TID has a similar candidate period limit as Apriori. The fascinating component is that it doesn't utilize a database for checking support after the chief pass. An encoding of the candidate itemsets used as a piece of the past pass is used. In later passes, the range of encoding can be a lot smaller than the database, saving examining effort.

### 4.2 *Proposed Apriori Hybrid Algorithm*

Considering the preliminary recognitions, the Apriori-Hybrid methodology was created which uses Apriori as an aspect of the basic passes and changes to Apriori-TID when it expects that the set  $l_k$  at the finish of the pass will fit in memory. Apriori has better execution in earlier passes, and Apriori-TID beats Apriori in later passes. This way, an assessment of  $l_k$  around the finish of each pass is crucial. Also, there is a cost commitment of changing from Apriori to Apriori-TID. The proposed algorithm is given in Fig. 1. This figuring relies upon the idea that it isn't essential to use a similar computation in all overlooks data.

```

for all large k-items  $l_k$ ,  $k \geq 2$  do
begin
     $H_1 = \{\text{consequents of rules derived from } l_k \text{ with one item in the}$ 
     $\text{consequent}\};$ 
call Hybrid-Rulegen ( $l_k$ ,  $H_1$ );
end

function Hybrid-Rulegen ( $l_k$ : large k-itemset,  $H_m$ : set of m-item consequents)

begin
if ( $k > m + 1$ ) then
begin
     $H_{m+1} = \text{apriori-gen}(H_m);$ 
for all  $h_{m+1} \in H_{m+1}$  do
begin
    conf = support( $l_k$ ) / support( $l_k - h_{m+1}$ )  $\Rightarrow h_{m+1};$ 
if (conf  $\geq$  minconf) then
begin
    output the rule  $(l_k - h_{m+1}) \Rightarrow h_{m+1};$ 
    confidence = conf; support = support( $l_k$ );
end
else
    delete  $h_{m+1}$  from  $H_{m+1};$ 
end
call Hybrid-Rulegen( $l_k$ ,  $h_{m+1}$ );
end
end
end

```

**Fig. 1** Proposed apriori hybrid algorithm

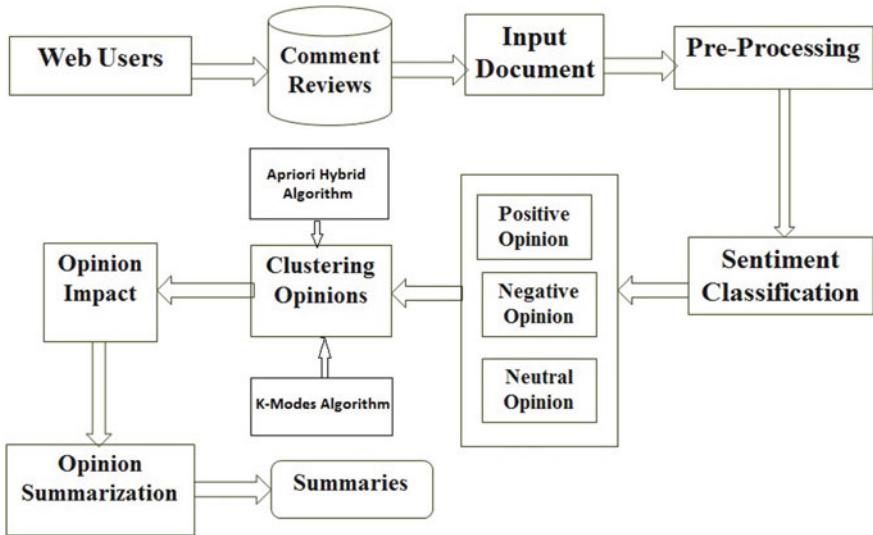
### 4.3 K-Modes Clustering

The k-modes calculation extends the k-implies worldview to bunch straight out information by utilizing (i) a basic coordinating divergence measure for downright questions, (ii) modes rather than means for groups and (iii) a recurrence based technique to redesign modes in the k-implies style to “minimize the grouping cost” capacity of bunching. Since the k-modes calculation utilizes “the same bunching process as k-means, it safeguards the productivity of the k-implies calculation”.

Let  $X$ ,  $Y$  be two unmitigated items portrayed by  $m$  all out characteristics. The aggregate characterizes the basic difference measure in the middle of  $X$  and  $Y$  befuddles of the relating characteristic estimations of the two articles. The littler the quantity of crisscrosses is the more comparative the two articles.

Formally,

$$d(X, Y) = \sum_{j=1}^m \delta(x_j, y_j) \quad (1)$$



**Fig. 2** Feature selection for clustering

where,

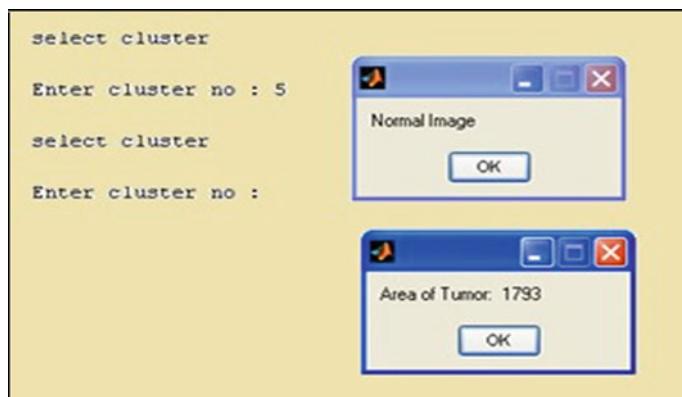
$$\delta(x_j, y_j) = \begin{cases} 0 & (x_j = y_j) \\ 1 & (x_j \neq y_j) \end{cases} \quad (2)$$

The architectural diagram is given in Fig. 2. The process of clustering with “feature selection involves the input variable (represent by full dataset features), and the final output variable is clustering pattern based on selected features from the previous feature selection process”.

## 5 Experimental Results

Results introduced for “our methodology depends on the mix of  $N = 100$  K-modes clusterings, a significant high incentive to guarantee that assembly of the strategy is guaranteed”. The detailed bunching blunder and the standard deviation is normal of 50 executions of the entire procedure. It very well may be seen that “the bunching results have improved with less standard deviation in blunder when our proposed strategy picked the modes in contrast with the irregular choice of introductory modes”.

The accompanying screen (Fig. 3) capture discloses whether the given information picture is typical or anomalous and finds the tumour’s territory.



**Fig. 3** Finding tumor

## 6 Conclusion

This paper gives an investigation of arranged specialized and audit papers on bosom tumour recognizable proof and visualization issues. It investigates that information preparing strategies supply a pleasant guarantee to reveal designs covered up in the data that encourage clinicians to make choices. From the above study, it is resolved that the precision for the finding investigation of various connected information preparing characterization strategies is extremely worthy and might encourage the restorative experts to choose for early recognizable proof and keep away from the biopsy. In this paper, we propose a sparing crossover model for forecast the inspiration and cynicism of the carcinoma depending upon the preparation information set. Various affiliation standard mining calculations are joined to support the precision inside the positive and negative forecast. An approach like this offers the specialists better choice some assistance with making by which valuable existence of a few patients can be spared.

## References

1. R.C. Gonzalez, R.E. Woods, *Digital Image Processing* (Prentice Hall, New Jersey, 2002)
2. D. Winston Paul, S. Balakrishnan, A. Velusamy, Rule based hybrid weighted Fuzzy classifier for tumor data. Int. J. Eng. Technol. (UAE) **7** (4.19), 104–108 (2018). <https://doi.org/10.14419/ijet.v7i4.19>
3. D. Delen, N. Patil, Knowledge extraction from prostate cancer data, in *The 39th Annual Hawaii International Conference on System Sciences*, pp. 1–10 (2006)
4. National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER) Program Public-Use Data (1973–2008). Cancer Statistics Branch (2011)
5. S. Guha, R. Rastogi, K. Shim, *CURE: An Efficient Clustering Algorithm for Large Databases*. Published in the Proceedings of the ACM SIGMOD Conference (1998)

6. S. Vasu, A.K. Puneeth Kumar, T. Sujeeth, S. Balakrishnan, A machine learning based approach for computer security. *J. Adv. Res. Dyn. Control Syst.* **10** (11-Special issue), 915–919 (2018)
7. Huang, *Extensions to the K-means Algorithm for Clustering Large Data Sets with Categorical Values, Data Mining and Knowledge Discovery*, Netherlands (1998), pp. 283–304
8. Y. Sun, Q. Zhu, Z. Chen, An iterative initial points refinement algorithm for categorical data clustering. *Pattern Recogn. Lett.* **23**, 875–884 (2002)
9. P.S. Bradley, U.M. Fayyad, Refining initial points for k-means clustering, in *Proceedings of the 15th International Conference on Machine Learning (ICML'98)*, SanFrancisco, Morgan Kaufmann (1998)
10. S.S. Khan, A. Ahmad, Computing initial points using density based multiscale data condensation for clustering categorical data, in *2nd International Conference on Applied Artificial Intelligence, ICAI'03*, Kolhapur, India (2003)
11. S. Kant, T.L. Rao, P.N. Sundaram, An Automatic and stable clustering algorithm. *Pattern Recogn. Lett.* **15**(6), 543–549 (1994)
12. A. Fred, A.K. Jain, Data clustering using evidence accumulation, in *Proceedings of the International Conference on Pattern Recognition (ICPR)*, Quebec City (Aug 2002)
13. A.L. Fred, Finding consistent clusters in data partitions, in *Multiple Classifier Systems, volume LNCS 2096*, ed. by J. Kittler, F. Roli (Springer), pp. 309–318 (2001)
14. Topchy, A., Jain, A.K., Punch, W. (2003), "Combining Multiple Weak Clusterings", in Proceedings of the IEEE International Conf. Data Mining, USA, pp. 331–338, 2003.
15. A. Sudha, Utilization of data mining approaches for prediction of life threatening diseases survivability. *Int. J. Comput. Appl.* **41** (17) (Mar 2012)
16. V. Chauraisa, S. Pal, Data mining approach to detect heart diseases. *Int. J. Adv. Comput. Sci. Inf. Technol. (IJACST)* **2**(4), 56–66 (2013)
17. V. Chauraisa, S. Pal, Early prediction of heart diseases using data mining techniques. *Carib. J. Sci. Tech.* **1**, 208–217 (2013)
18. S. Balakrishnan, J. Janet, K. Sujatha, S. Rani, An efficient and complete automatic system for detecting lung module. *Indian J. Sci. Technol.* **11** (26). <https://doi.org/10.17485/ijst/2018/v11i26/130559>
19. K. Sujatha, S. Krishnan, S. Rani, M. Bhuvana, Biometric authentication system with hand vein features using morphological processing. *Indian J. Sci. Technol.* **11** (26) (2018). <https://doi.org/10.17485/ijst/2018/v11i26/130558>
20. P. Palanikumar, S. Geofrin Shirly, S. Balakrishnan, An effective two way classification of breast cancer images. *Int. J. Appl. Eng. Res.* **10** (21) (2015), pp. 42472–42475. ISSN 0973-4562
21. D. Dharunya Santhosh, S. Balakrishnan, C.K. Dhivyasshri, N. Prakash, Biometric authentication system using matlab. *Int. J. Eng. Technol. (UAE)* **7** (4.19), 101–103 (2018)
22. B.S. Balaji, S. Balakrishnan, K. Venkatachalam et al., Automated query classification based web service similarity technique using machine learning. *J. Ambient Intell. Human Comput.* (2020). <https://doi.org/10.1007/s12652-020-02186-6>
23. S. Balakrishnan, K. Aravind, A. Jebaraj Ratnakumar, A novel approach for tumor image set classification based on multi-manifold deep metric learning. *Int. J. Pure Appl. Math.* **119** (10c), 553–562 (2018)
24. S. Balakrishnan, D. Deva, Machine intelligence challenges in military robotic control. *CSI Commun. Mag.* **41** (10), 35–36 (Jan 2018)
25. J. Rathnakumar, S. Balakrishnan, Machine learning based grape leaf disease detection. *J. Adv. Res. Dyn. Control Syst.* **10** (08-Special Issue), 775–780 (2018)
26. D. Prabha, R. Siva Subramanian, S. Balakrishnan, M. Karpagam, Performance evaluation of Naive Bayes classifier with and without filter based feature selection. *Int. J. Innov. Technol. Exploring Eng. (IJITEE)* **8** (10), 2154–2158. ISSN: 2278-3075 (Aug 2019)
27. W. He, A. Juette, E.R.E. Denton, A. Oliver, R. Martí, R. Zwiggelaar, *A Review on Automatic Mammographic Density and Parenchymal Segmentation*, vol. 2015, Article ID 276217 (Hindawi Publishing Corporation) (*Int. J. Breast Cancer*) (2015), 31 pp

# Deep Learning Approach Based on ADASYN for Detection of Web Attacks in the CICIDS2017 Dataset



Kuljeet Singh, Amit Mahajan, and Vibhakar Mansotra

**Abstract** The increased usage of web-based applications, vast number of web users, and emphasis on data storage on the web have been both productive and constructive but it also exposes vulnerabilities in the system. These applications and the confidential data associated with them and stored on the web need to be protected from the growing number of web-based attacks. The contribution of this paper is classifying the three most common web-attacks namely Brute force, SQL injection (SQLI), and cross-site scripting (XSS) using deep learning (DL). ADASYN, a variant of Synthetic Minority Oversampling Technique (SMOTE) has been employed to address the imbalance situation as these attacks have fewer training instances in the CICIDS2017 dataset. Experimental results have shown improvement in the detection of these attacks using the proposed approach.

**Keywords** SQL injection · Cross site scripting · SMOTE · ADASYN · Web-attacks · CICIDS2017

## 1 Introduction

In the ubiquitously connected world, web-based applications are getting more prevalent with each passing day. Many new applications are created daily which are web-based and several standalone applications are getting converted to web-oriented applications. Moreover, techniques like cloud computing and edge computing have encouraged more data storage on the web. This data-driven approach of web applications has led to data increasing in terms of size and becoming more valuable as confidential data including user names, passwords, monetary and financial information stored in the web servers. This enables web applications to get more productive and convenient while also exposing them to security vulnerabilities thus becoming more prone to cyber-attacks. These web applications along with websites and web users must be protected and safeguarded from the cyber-attacks.

---

K. Singh (✉) · A. Mahajan · V. Mansotra

Department of Computer Science & IT, University of Jammu, J&K, Jammu, India  
e-mail: [khalsakuljeet004@gmail.com](mailto:khalsakuljeet004@gmail.com)

Some of the common attacks occurring on web-based applications are Brute force, SQLI, and XSS. The objectives behind these attacks might be disrupting website functioning, getting unauthorized access to the database, or stealing user credentials. Brute force is a simple yet popular method for retrieving the login credentials or performing other tasks. SQLI attack is launched to gain entrance into the database and it occurs when data from the client-side in the form of SQL statements are entered in an unsafe manner [1]. There are different types of SQLI attacks like Tautologies, Logically Incorrect queries, Union Query, and Piggy-backed queries among others. The attacker's aim behind these attacks is to bypass authentication and extract information from the database; execute remote commands; and launch other attacks like Denial of Service (DoS). XSS is another type of attack which occurs by inserting malicious code into a legitimate web application which upon visited by the user results in the infectious code getting into the user system [2]. This attack affects the user utilizing the web service and not the web application itself.

In this paper, deep learning is used to classify web-attacks attempts in the CICIDS2017 dataset. ADASYN, a variant of SMOTE is considered for oversampling of the minority class labels. There are fewer datasets that contain records of web-based attacks and those datasets that actually contain records have insufficient training instances which led to classifiers performing poorly and showing bias towards the majority class. So, resampling is done to minimize this problem and achieve better results not for an overall model but for the minority class also.

The rest of the paper is categorized as: Sect. 2 discusses the related work in detecting web-attacks and Sect. 3 briefly describes the material used and technique employed. Section 4 contains the methodology of the proposed work and Sect. 5 explains the results of the experimentation process while Sect. 6 concludes the paper.

## 2 Related Work

Numerous detection solutions for web-based attacks have been found in the literature. This section briefly discusses the work done in the field.

Nagpal et al. performed a survey regarding the background of SQLI attacks, discussing various vulnerabilities occurring and the different types of SQLI attacks performed on the web applications [3]. The authors in their survey discussed various mitigation techniques ranging from encryption, binding, and obfuscation to defensive coding. Review surrounding query translation, cryptography, Machine learning (ML), and other detection strategies has been performed. Hu et al. also presented an extensive review of SQLI attacks, the impact of SQLI on web applications, different types of SQLI attacks deployed by hackers, and several criteria used for detecting and preventing attacks [4]. The authors also discussed the various approaches and architectures proposed by the researchers for detecting and mitigating SQLI attacks. Kaur et al. used ML techniques for detecting and analyzing Blind XSS which is a subset of stored XSS and is more lethal compared to normal XSS [5]. Given the nature of these attacks, it is difficult to detect them by identifying malicious payloads,

so behavior analysis using ML methods was done. For this purpose, linear Support Vector Machine (SVM) was deployed and its classification helps in determining distinct features between Blind and normal XSS.

Betarte et al. performed characterization of the problem and define unique scenarios depending upon the availability of attack data for training purposes [6]. Their work revolves around two approaches; first when normal and attack data are both available; second when no attack data is available and only valid data is available for training. The former approach utilizes a multi-class solution while the latter contend with only a single-class solution. Lin et al. applied the Convolutional Neural Networks on KDD 99 dataset to classify records as normal or abnormal [7]. They developed a behavior-based classifier learning model by training a CNN to extract the behavior features (32 out of 41 are selected) and identify the class of a record by using collected statistical data. Much emphasis has been laid on attacks with higher frequencies as top-six attack types are selected for detailed analysis.

For detection of web attacks, moh et al. proposed multi-stage log analysis [8]. The core of their architecture revolves around ML and pattern matching with Bayes Net and Kibana as respective algorithms for these two. Model evaluation of 2-stage architecture was performed on 10,000 logs in a series of experiments. Parsaei et al. addressed the problem of class imbalance by using a hybrid approach to oversample the minority class labels [9]. The dataset used by the authors in their experiments is the NSL-KDD which contains two attacks, U2R and R2L, suffering from imbalance situation owing to their lesser training instances. SMOTE and cluster center and nearest neighbor (CANN) are part of the hybrid approach while leave one out method (LOO) is considered for selecting relevant features. Maseer et al. have performed a comparative analysis of 10 common ML/DL techniques for detecting web attacks in the CICIDS2017 dataset [10]. The employed techniques include supervised algorithms like ANN, SVM, KNN, DT, CNN, RF, and NB unsupervised techniques such as k-means, expectation maxim and SOM. Experimental results have shown that the NB, KNN and DT have outperformed the other models as the majority of the models have failed in the proper detection of SQLI and XSS attacks. Hoang et al. proposed a detection model which comprises DT, RF and SVM as classification techniques, and the experiments were conducted on HTTP Param dataset and also on real web logs [11]. Their web attacks detection model was capable of detecting SQLI, XSS, command injection and path traversal type of attacks.

### 3 Material and Methods

#### 3.1 Dataset

The dataset used in this study is CICIDS2017 which contains records of various attack types [12]. The focus of our work is the web attacks for which the sub-file of the dataset named Thursday-WorkingHours-MorningWebAttacks.pcap\_ISCX.csv is

**Table 1** Number of instances for each label

Label	Instances
BENIGN	168,186
Web attack-Brute force	1507
Web attack-XSS	652
Web attack-Sql injection	21

considered. Each record contains 78 statistical features and the number of instances for each label is shown in Table 1.

### 3.2 ADASYN

SMOTE is an oversampling technique used for balancing datasets by creating new synthetic instances from existing insufficient instances [13]. This is a type of data augmentation where minority class records are used to create new instances. The basic working of SMOTE is shown in the following steps.

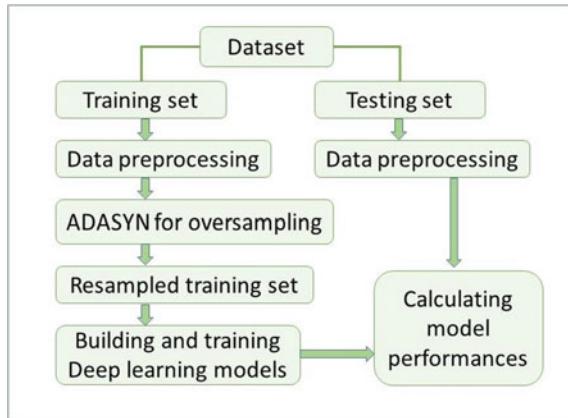
- (a) Select any random individual record in the minority class.
- (b) Find the nearest neighbors of each individual record by calculating the Euclidean distance between that record and other records in the particular class.
- (c) Based on the sampling rate, nearest neighbors are chosen and new synthetic records are created.

ADASYN is a variant or extension of SMOTE also generates synthetic instances or samples but in selective regions [14]. ADASYN uses density distribution as a criterion for sampling as the regions in the feature space where the density of minority class is low are only selected for synthetic generations of instances. To compensate for the skewed distributions, it adaptively changes the weights of the different minority samples depending upon the region in the feature space. These are the regions that often result in poor performance when classified by the ML/DL models. This is the key difference with SMOTE as SMOTE treats all instances in minority class as equal but in ADASYN, the instances with difficulty in learning are given more importance. ADASYN shifts the decision boundary of the classifier used which helps in focusing tricky instances thereby contributing in improved model performance.

## 4 Research Methodology

The overall methodology is shown in Fig. 1. Firstly, the dataset is partitioned into 80/20 train and test ratios. Data preprocessing is then performed and applied to both

**Fig. 1** Flow of the proposed system



sets and it includes handling missing values and normalizing the data using standardization. Then ADASYN is applied to the training instances which synthetically increases the number of minority training instances resulting in the resampled dataset shape for the training set. Different values for sampling ratios are experimented with resulting in different training instances for the minority class labels. These different sampling rates enable different training sets for model training with a unique number of training instances in the minority class. ADASYN employed alters the shape of the training set owing to sampling rate but it has no effect upon test data.

After the dataset has been resampled, the training instances are fed to the model built on CNN using 1d-convolution. The architecture of the model used in the experiment consists of an input layer followed by two convolution layers each immediately followed by a dropout layer which is then flattened to be connected to the dense layer which is finally attached to the output layer. Fine-tuning of all the parameters and hyper-parameters are done to achieve optimal performance of the CNN model built. Comparative analysis of all the models built is done upon testing data.

## 5 Experimental Results and Analysis

### 5.1 Evaluation Indicators

Various metrics used in classification problems are considered keeping in mind the imbalance nature of the dataset. As accuracy is not considered good indicator of model performance in the imbalance situations, it is not included as one of the evaluation metrics. Several metrics used in the experimental analysis are:

- (a) TP: Correct label classified as correct.
- (b) FN: Correct label classified as incorrect.

- (c) FP: incorrect label classified as correct.
- (d) Precision: Ratio of number of correctly classified labels to the number of all labels classified as correct.
- (e) Recall: Ratio of number of correctly classified labels to the actual number of labels.
- (f) F1\_sc: considered as harmonic mean of precision and recall.

## 5.2 Results

Before actual model training, ADASYN is applied to the training instances to increase the number of samples. Here the thing to consider is how much the samples must be increased. Too much increase will deviate the training set from reality and a minor increase will result in no change in performance of the model. So, different sampling ratios are experimented with resulting in different training instances for each minority class. This results in five cases with their sampling strategy discussed below and number of training samples for each class is shown in the Table 2.

- Case 1** No sampling is done and training instances are unchanged after 80/20 split.
- Case 2** All minority class labels have their instances increased by a multiplier of 20.
- Case 3** Brute force attack is kept unchanged while instances of XSS and SQLI attacks are increased by a multiplier of 20.
- Case 4** Instances of brute force attack are increased by a multiplier of 10 while other two increased 20 times.

**Table 2** Number of training instances for different sampling ratios

	Sampling strategy	BENIGN: 0	Web attack-Brute force:1	Web attack-XSS:2	Web attack-Sql injection:3	Training shape
Case 1	No sampling	134,538	1214	524	16	(136,292, 78)
Case 2	1,2,3: *20	134,538	24,186	10,558	318	(169,618, 78)
Case 3	2,3: *20	134,538	1214	10,558	318	(146,628, 78)
Case 4	1: *10; 2,3: *20	134,538	12,162	10,558	318	(157,576, 78)
Case 5	1,2:6000; 3: *20	134,538	5906	5983	318	(146,745, 78)
Case 6	1:6000; 2:7500; 3: *20	134,538	5916	7508	318	(148,280, 78)

**Case 5** Manual selection for a number of instances done as both brute-force and XSS fixed to 6000 instances while SQLI increased like previous cases.

**Case 6** Similar to case5 with XSS increased to 7500 instances.

It is to be noted that despite specifying the number of instances, the actual synthetic samples generated by ADASYN might be slightly more or lesser. For instance, in case5 labels 1and 2 (brute force and XSS) have been set to 6000 resampled instances but the samples generated by ADASYN are 5906 and 5983 respectively. Results for each class labels using different cases of sampling strategy are shown in Table 3.

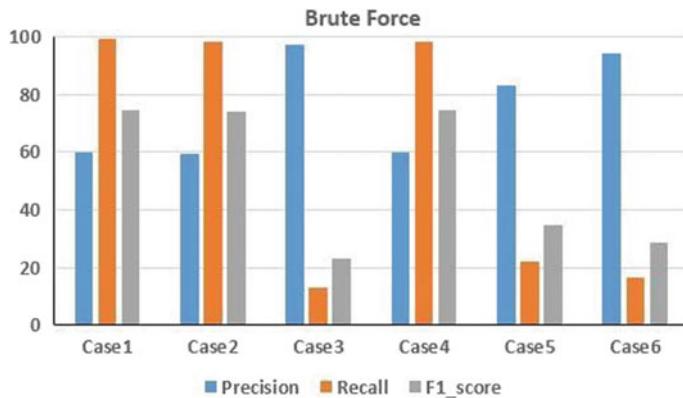
From the results, it can be analyzed that the XSS attack was not getting detected by the model, as it has very less TP, before sampling was applied. After sampling with different ratios, detection of XSS has improved significantly but it has come at cost of a Brute force attack. A trade-off has been observed between these two as increasing the training instances of one during sampling led to poor detection of the

**Table 3** Performance metrics for different cases

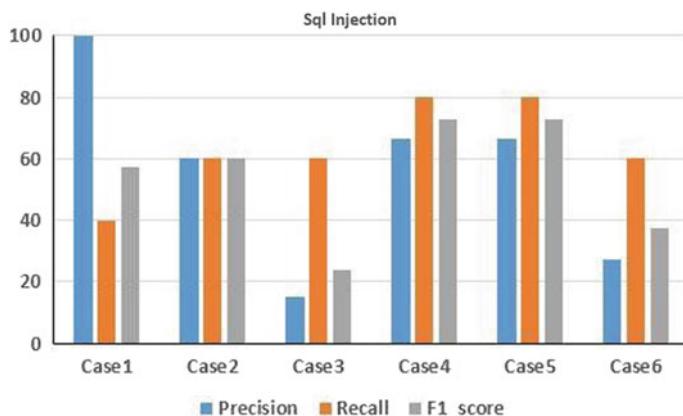
	Labels	TP	FN	FP	Precision	Recall	F1_sc
Case 1	BENIGN	33,576	72	7	99.98	99.79	99.88
	Brute force	291	2	195	59.88	99.32	74.71
	Sql injection	2	3	0	100.00	40.00	57.14
	XSS	2	126	1	66.67	1.56	3.05
Case 2	BENIGN	33,575	73	6	99.98	99.78	99.88
	Brute force	289	4	195	59.71	98.63	74.39
	Sql injection	3	2	2	60.00	60.00	60.00
	XSS	1	127	3	25.00	0.78	1.52
Case 3	BENIGN	33,574	74	6	99.98	99.78	99.88
	Brute force	38	255	1	97.44	12.97	22.89
	Sql injection	3	2	17	15.00	60.00	24.00
	XSS	122	6	313	28.05	95.31	43.34
Case 4	BENIGN	33,574	74	6	99.98	99.78	99.88
	Brute force	289	4	193	59.96	98.63	74.58
	Sql injection	4	1	2	66.67	80.00	72.73
	XSS	1	127	5	16.67	0.78	1.49
Case 5	BENIGN	33,575	73	5	99.99	99.78	99.88
	Brute Force	64	229	13	83.12	21.84	34.59
	Sql Injection	4	1	2	66.67	80.00	72.73
	XSS	118	10	293	28.71	92.19	43.78
Case 6	BENIGN	33,575	73	5	99.99	99.78	99.88
	Brute force	49	244	3	94.23	16.72	28.41
	Sql injection	3	2	8	27.27	60.00	37.50
	XSS	122	6	309	28.31	95.31	43.65

other during classification. Figures 2, 3 and 4 displays the graph chart representing performance evaluation metrics for the Brute force, Sql Injection and XSS attack labels respectively.

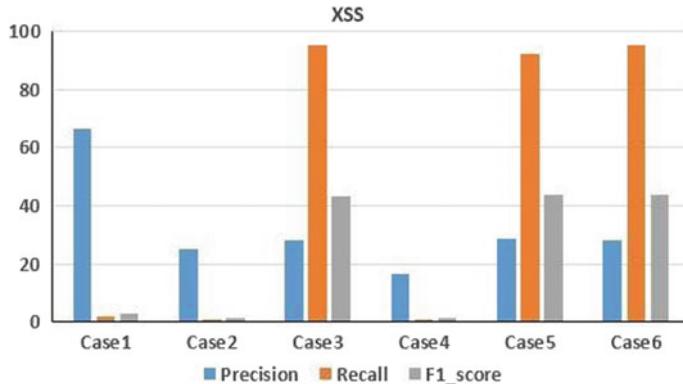
Analyzing Figs. 2 and 4, model performance for detection of XSS has improved in the latter cases, which involves selective sampling or more sampling instances for XSS, compared to Brute force, whose detection rate is better with no sampling but decreases in the latter cases of sampling.



**Fig. 2** Performance metrics for the Brute force



**Fig. 3** Performance metrics for the Sql injection



**Fig. 4** Performance metrics for the XSS

## 6 Conclusion

The study proposed a DL model for identifying Web-attacks instances in the CICIDS2017. The dataset used suffered from an imbalance situation which affects the performance of the model in detecting minority class labels. ADASYN, a variant of SMOTE was used for resampling of training instances in the minority class. Different values for the sampling ratio were experimented with which leads to different sets of training data. A classification model based on CNN has been built which is trained using different training sets and consequently, evaluation and comparative analysis have been done. Experimental results have shown improvement in the detection of web attacks after resampling has been done.

## References

1. W.G. Halfond, J. Viegas, A. Orso, A classification of SQL-injection attacks and countermeasures, in *Proceedings of the IEEE International Symposium on Secure Software Engineering*, vol. 1 (IEEE, Mar, 2006), pp. 13–15
2. B.B. Gupta, P. Chaudhary, *Cross-site Scripting Attacks: Classification, Attack, and Countermeasures* (CRC Press, 2020)
3. B. Nagpal, N. Chauhan, N. Singh, A survey on the detection of SQL injection attacks and their countermeasures. *J. Inf. Process. Syst.* **13**(4), 689–702 (2017)
4. J. Hu, W. Zhao, Y. Cui, A survey on SQL injection attacks, detection and prevention, in *Proceedings of the 2020 12th International Conference on Machine Learning and Computing* (Feb, 2020), pp. 483–488
5. G. Kaur, Y. Malik, H. Samuel, F. Jaafar, Detecting blind cross-site scripting attacks using machine learning, in *Proceedings of the 2018 International Conference on Signal Processing and Machine Learning* (Nov, 2018), pp. 22–25
6. G. Betarte, Á. Pardo, R. Martínez, Web application attacks detection using machine learning techniques, in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (IEEE, Dec 2018), pp. 1065–1072

7. W.H. Lin, H.C. Lin, P. Wang, B.H. Wu, J.Y. Tsai, Using convolutional neural networks to network intrusion detection for cyber threats, in *2018 IEEE International Conference on Applied System Invention (ICASI)* (IEEE, Apr, 2018), pp. 1107–1110
8. M. Moh, S. Pininti, S. Doddapaneni, T.S. Moh, Detecting web attacks using multi-stage log analysis. in *2016 IEEE 6th International Conference on Advanced Computing (IACC)* (IEEE, Feb, 2016), pp. 733–738
9. M.R. Parsaei, S.M. Rostami, R. Javidan, A hybrid data mining approach for intrusion detection on imbalanced NSL-KDD dataset. *Int. J. Adv. Comput. Sci. Appl.* **7**(6), 2025 (2016)
10. Z.K. Maseer, R. Yusof, N. Bahaman, S.A. Mostafa, C.F.M. Foozy, Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset. *IEEE Access* **9**, 22351–22370 (2021)
11. X.D. Hoang, T.H. Nguyen, Detecting common web attacks based on supervised machine learning using web logs. *J. Theor. Appl. Inf. Technol.* **99** (6) (2021)
12. I. Sharafaldin, A.H. Lashkari, A.A. Ghorbani, Toward generating a new intrusion detection dataset and intrusion traffic characterization, in *4th International Conference on Information Systems Security and Privacy, ICISSP* (2018), pp. 108–116
13. N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority oversampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002)
14. H. He, Y. Bai, E.A. Garcia, S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)* (IEEE, June, 2008), pp. 1322–1328

# FUNDUS and OCT Image Classification Using DL Techniques



Akanksha Bali and Vibhakar Mansotra

**Abstract** The phrase “Medical Diagnosis” alludes to the means of identifying a person’s disease symptoms and signs. Decision-making in medical diagnosis is a challenging task as it requires highly specialized medical experts for interpreting the parameters and symptoms of the diseases. Different types of studies such as pathology, ophthalmology, oncology, etc. are conducted to understand the propagation of different kinds of diseases. Ophthalmology is the branch that deals with the study of the eye. The Eye is a sensitive and vital sense organ of the human body that responds to light and allows vision. For living a quality life, a good vision is a necessity of human beings. In FUNDUS and optical coherence Tomography (OCT) pictures, DL (DL) have excelled. The majority of prior research has concentrated on recognizing a specific FUNDUS and OCT illness. This work reviews various papers on the DL technique for FUNDUS and OCT classification. Several classification approaches had been communicated in the literature for the automatic classification of FUNDUS and OCT images in which DL techniques outperformed. Different DL (DL) techniques for automatic eye diseases classification had been discussed in this paper and results are compared on the basis of accuracy, F1-score, and AUC.

**Keywords** Convolution neural network (CNN) · DL (DL) · Ensemble learning (EL) · FUNDUS images · OCT images · Multi class classification · Multi label classification · And transfer learning (TL)

## 1 Introduction

Diseases of the FUNDUS can cause vision loss and blindness. Diabetes retinopathy, cataract, and age-related macular degeneration (AMD) are common eye diseases that affect vision. AMD is the first blinding eye disease in the world. Today’s World demands automated and accurate diagnostic processes. With the emergence of DL

---

A. Bali · V. Mansotra

Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, Jammu 180006, India

e-mail: [akankshabali5@gmail.com](mailto:akankshabali5@gmail.com)

and image processing techniques, there has been substantial progress in the automation of the detection/prediction process for various eye diseases. Technology intervention has provided easy access and gained good popularity in the early detection of the disease. DL models for computer vision have rapidly evolved. Convolutional neural networks can automatically learn image features. They've done well in image classification and target detection. Their representation ability compensates for traditional image extraction flaws. The novel procedure has also been applied in the medical areas to improve the management of various health problems. Diabetic retinopathy, AMD, and glaucoma were the illnesses that were investigated the most.

There are many types of imaging available for eye diseases. But in this paper, only FUNDUS and OCT images are considered. FUNDUS images are taken through a FUNDUS imaging camera in which serial photographs of the interior of our eyes were taken via pupil and also used to investigate retinal, lens, and optic disc structures. On the other hand, OCT images are taken through lightwaves in which cross-sectional images of the retina are generated so that their thickness is easily calculated. As FUNDUS imaging (photography) is non-invasive, OCT imaging is invasive in form. As there are three types of classification, binary class (two classes 0 or 1 with binary cross-entropy), multi-class (2 or more classes with softmax but one disease classified at a time), multilabel (more than one disease occurred simultaneously), there are three types of datasets available in the database such as binary class data, multi-class data, and multilabel data.

## 2 Related Work

Various DL methods had been discussed in the literature that had shown remarkable outcomes in the field of automatic classification of FUNDUS and OCT images. The DL models give favorable results in classifying eye diseases while working on huge amount of datasets and much accurate than expert results. In the field of medical FUNDUS and OCT images, no of researches for automatic classification have been done by different researchers based on DL techniques. This motivated to create review on DL algorithms to find classification between various FUNDUS and OCT diseases.

### 2.1 OCT Image Classification

Kermany et al. proposed a DL tool for detecting treatable blinding retinal diseases. Mendeley trained on CNV: 37,206, DME: 11,349, DRUSEN: 8617, and Normal 51,140 normal, respectively [1]. In this paper [2], Kaymak et al. used alexnet model on the Kermany dataset for AMD and DME detection. In this paper [3], Das et al. introduced CNN based on multi-scale deep feature fusion for OCT images classification on the Kermany dataset. Bhowmik et al. used a kaggle subset of the Kermany

dataset to speed up training. This data collection trains the intense learning algorithms on DRUSEN, NORMAL, and CNV pictures to employ AI to treat aging, and macular degeneration [4]. In this paper [5], Li et al. presented automated detection of retinal diseases using vgg16 on the Mendeley dataset. In this paper [6], Alqudah proposed SD-OCT using proposed automated CNN for\_multiclass\_problem. In this paper, \_Mendeley\_dataset\_as\_well\_as\_additional\_AMD\_of\_900\_images\_collected\_from\_duke dataset. Also, additional 1250 images were collected from different hospitals for testing purposes. Fang et al. suggested IFCNN-based retinal OCT image classification. The dataset was collected from Shiley Eye Institute for OCT B-scans from 4686 patients. A total of a training dataset of 83,484 images (11,348 DME, 37,205 CNV, 8616 DRUSEN, and 26,315 Normal images), and 1000 test images (250 in each category) were used [7]. In this paper [8], Yoo et al. illustrated FSL using a cyclegan that enhances the efficiency of DL in the rare OCT eye diseases diagnosis using inceptionv3 architecture on Mendeley dataset and custom retinal image datasets (30 macular hole, 30 central serous chorioretinopathy, 19 Stargardt disease, 24 macular telangiectasia, and 16 retinitis pigmentosa) collected from Google search engine, applied augmentation method of CycleGAN, and generates 3000 augmented images per class. Rajagopalan et al. proposed a CNN model for retinal disorders diagnosis using Mendeley dataset of 12,000 images (3000 in each class) in which training images used were 8000 and testing images used were 4000 (1000 per class) [9]. For AMD identification, Thomas et al. suggested a new multipath convolutional neural network design. Training had been done on Mendeley and assessed on four datasets named OCTID, Mendeley, SDOCT noor, and Duke dataset. The testing dataset contains 250 each for dry AMD images, wet AMD images, and normal images. The model had achieved an overall accuracy of 99.61% on OCTID, 96.67% on Duke, and 93.87% on SDOCT [10].

## 2.2 *FUNDUS Image Classification*

### 2.2.1 **Binary and Multiclass Classification**

Ghosh et al. proposed automatic DR detection using CNN and data augmentation through brightness, rotation on the kaggle2 dataset [11]. Raju et al. used custom CNN, augmentation method through rotation, translation, zoom, flip, color augmentation, image centering on five class eyepacs kaggle2 dataset for DR detection and its stages [12]. Wan et al. proposed Deep CNN for DR classification by image-based approach. Kaggle2 dataset of 35,126 color FUNDUS images (five stages or classes, namely 25,810 normal, 2443 mild, 5292 moderate, 873 severe, and 708 PDR) were collected [13]. Chen et al. used inceptionv3 and image augmentation using flipping, cropping, rotation, horizontally, and vertically scaling for DR detection [14]. Lin et al. used custom CNN, an augmentation that leads to 33,000 images through rotation and flipping on a 5-class eyepacs dataset for comparative evaluation of severe DR [15].

Raghavendra et al. presented an 18 CNN layer and LDA for glaucoma detection on a custom image dataset of 1426 images (589 N, 837 G) collected from Kasturba college, Manipal [16]. Serener et al. proposed an early and advanced glaucoma detection technique using augmentation and ResNet50 and googlenet on a custom image dataset of 1544 FUNDUS images (788 Normal, 289 early G, and 467 advanced G) and RIM-ONE dataset of 158 images (118 normal, 12 early, 14 moderate, 14 advanced glaucoma category) [17]. In this paper [18], Sun et al. proposed UDAGAN and RESNET based methods for glaucoma diagnosis on DRISHTI-GS, REFUGEE, and private dataset and achieved accuracies of 75.0, 89.75, and 82.0 respectively. The private dataset1 contains 400 images (63 Glaucoma, 337 Non-Glaucoma) and the private dataset2 contains 9273 images (1089 Glaucoma, 8184 non Glaucoma). In this paper [19], Rehman et al. used ensemble classifier (AlexNet, InceptionV3, InceptionResNetV2, and NasNet-Large) and single classifier NasNET method on ACRIMA dataset of 705 images (309 N, 396 G), ORIGA dataset of 650 images (482 N, 168 G), RIM-ONE of 455 images (261 N, 194 g), AFIO of 124 images (85 N, 39 g), HMC of 55 images (40 n, 15 g).

Grassmann et al. outlined AMD classification algorithm development depending on a large-scale collection of FUNDUS images that were cropped, rotated, adjusted aspect ratio. The dataset used were AREDS containing 120,656 images (training: 86,770, testing: 12,019, validation: 21,867), and KORA dataset (5555 images) [20]. Tan et al. put forwarded an automated technique for AMD detection with FUNDUS images using fourteen layer CNN, ten-fold cross-validation on Kasturba dataset (402 normal and 708 AMD) [21]. Peng et al. evolved DeepSeeNet for automatic classification of AREDS patients (59,302 images) [22]. Heo et al. developed a DL method, and also used image data augmentation methods such as moving up, down, left, right, flipping, rotating, and zooming, 5-Fold Cross-Validation methods [23].

### 2.2.2 Multilabel Classification

Islam et al. suggested a classification model using contrast limited adaptive histogram equalization and CNN for eight ocular disease classifications [24]. He et al. proposed a CNN-based multi-label ocular disease classification depending on patient level. The 5000 patients' left and right eye images were analyzed for disease and label data. ResNet -101 produces better results 0.827 with the spatial correlation module [25]. In this paper [26], Gour et al. proposed an ophthalmology disease classification using VGG-16 transfer learning on two models firstly on the merged left and right eye image, and secondly on the separate left and right eye and then combined for classification. The model had achieved overall validation accuracy of 0.89. Wang et al. proposed a multilabel classification ensemble model (EfficientNetB3, InceptionResNetV2) on ODIR-19 FUNDUS images to directly detect multilabel eye diseases. Ten thousand images of left and right eyes were collected from 5000 patients having eight classes (1624 Normal, 1620 DR, 238 AMD, 305 Glaucoma, 308 Cataract, 149 Hypertension, 243 Myopia, 1393 others) were used. Also, 40 custom FUNDUS images were collected from <https://drive.grand-challenge.org/>. The model had achieved an

accuracy of 0.92, Final\_Score of 0.72, and FB\_score of 0.92 on image size 448\*448 using EfficientNetB3. [27]. Li et al. proposed multi-disease detection by using visual geometry group-16, residual neural network (ResNet-18, ResNet-50, ResNeXt-50, SE-ResNet50, SE-ResNeXt-50), Inception-v4, CaffeNet, and Densenet from international ocular disease intelligent recognition competition Peking university (ODIR 2019) dataset using DL identification and classification of eye disease [28]. In this paper [29], Cen et al. utilized the DCNN approach on 249,620 images having 275,543 labels obtained from various sources. In this paper [30], Bali et al. used one versus rest classifier and VGG approach for multiclass multilabel diseases prediction on ODIR-19 dataset and achieved validation accuracy of 90.85% and F1-score of 0.91. No fusion of left and right eye images was considered. The validation accuracy achieved in this paper outperforms the previous results obtained by different researchers in the ODIR-19 dataset.

Furthermore, multi-label disorders frequently impact OCT and FUNDUS, which are not visible to the human eye. As a result, the demand for disease detection is increasing.

It can be observed, from the literature survey:

- The absence of a completely automated system to do multi-label classification at a medically permissible rate.
- Research is mostly done by employing one or multiclass eye disease classification.
- The lack of a large number of multi-label datasets reduces the option of DL.

### 3 Comparative Study on OCT and FUNDUS Image Classification

This section compares FUNDUS and OCT images based on the dataset, the algorithm used, and its performance metrics (Accuracy, F1-score, and AUC). Different DL-based approaches for the automatic classification of eye OCT images and FUNDUS images introduced by different researchers had been discussed and outlined in the tabular form (Tables 1, 2, 3, 4, and 5). Table 1 shows the comparison between all the parameters of OCT image classification on the Kermany and Mendeley dataset. Figure 1 represents the bar-graph accuracy comparative analysis of OCT image classification on the Kermany and Mendeley dataset. Table 2 shows the comparison between all the parameters of FUNDUS image classification on the kaggle2 DR staging dataset. Figure 2 represents the bar-graph accuracy comparative analysis of FUNDUS image classification on the kaggle2 DR staging dataset. Table 3 shows the comparison between all the parameters of FUNDUS image classification on Glaucoma. Figure 3 represents the bar-graph accuracy comparative analysis of FUNDUS image classification on Glaucoma.

Tables 3 and 4 show the comparison between all the parameters of FUNDUS image classification on Glaucoma and AMD respectively. Figures 3 and 4 represent the bar

**Table 1** Accuracy comparison on OCT image dataset

Ref. No	Techniques used	Results	Remarks
[1]	Transfer learning	Overall training accuracy: 96.6%; Testing accuracy: 93.4%	Possible to attain a high accuracy with this four-class eye illness classifier using transfer learning technique on Kermany dataset
[2]	Alexnet architecture	Overall testing accuracy 97.1%	Alexnet_model_outperforms_transfer_learning_method_of Kermany et al. [1]
[3]	CNN based multiscale deep feature fusion classification algorithm	Overall testing accuracy 99.6%; Overall validation accuracy: 95.5%	The results yielded competitive results for DME, CNV, Drusen, normal detection and are comparable to advanced systems The results indicate that the proposed system surpasses the present transfer learning-based method in terms of accuracy on the Kermany dataset
[4]	vgg16 with inception v3	Overall testing accuracy 94%; Overall training accuracy: 99.94%	Highlights the applicability and benefits of using transfer learning in CNN models to acquire the most out of advanced models without the need for expensive hardware on the Kermany dataset
[5]	VGG16	Overall Testing Accuracy 98.6%	The proposed approach demonstrated higher competence in retinal OCT image detection on Mendeley Dataset
[6]	Proposed CNN	Testing accuracy: 95.30% for the 5-class; 98.56% for the 4-class	Probable to achieve high accuracy with this five class eye illness classifier using the proposed CNN technique on Mendeley Dataset
[7]	Iterative fusion CNN	Overall accuracy: 93.65%	Capable of predicting OCT disorders, which can speed up the diagnosis process and improve visual outcomes on Mendeley dataset

(continued)

**Table 1** (continued)

Ref. No	Techniques used	Results	Remarks
[8]	Inception v3 architecture, CycleGAN, finetuning, FiveFold cross validation	Without CycleGAN, overall accuracy: 93.9% With CycleGAN, overall accuracy: 92.1%	Without_cyclegan_classification_model_giving better performance_than_with_cyclegan on Mendeley Dataset
[9]	Proposed CNN	Testing accuracy: 97.01%	Utilized in real-time to attain fewer complications and more infrequent learnable metrics on Mendeley dataset
[10]	Multipath CNN	It achieved accuracies of 99.60% on Mendeley dataset	The results generated competitive results for dry AMD, wet AMD, normal detection and are equivalent to advanced systems. Multipath CNN beats other models in terms of accuracy

**Table 2** Accuracy comparison of multiclass Kaggle 2 fundus image dataset

Ref. No	Techniques used	Results	Remarks
[11]	CNN; Data augmentation	Accuracy: 95 and 85% for the 2-class and 5 class classification model respectively	Better performance with 2-class classification model than 5-class classification model
[12]	Custom CNN; Augmentation	Accuracy: 93.28%	In this research, DLA performs well in detecting DR staging in retinal imaging under a variety of imaging techniques
[13]	AlexNet, VGGNet, GoogleNet, ResNet; Data augmentation;	VGGNET-s best Accuracy: 95.68, AUC 0.9786	VGGNET outperforms other models in terms of accuracy
[14]	Inceptionv3; image Augmentation	Accuracy: 80%; Kappa Score: 0.64	The proposed method is memory efficient and gives less accuracy percentage
[15]	Custom CNN; augmentation	Accuracy: 86.10%; AUC: 0.92	Better performance with less GPU utilization

**Table 3** Accuracy comparison on glaucoma fundus image dataset

Ref. No	Techniques used	Results	Remarks
[16]	CNN (18 layer); LDA	Accuracy: 98.13	Better performance with less GPU utilization and less amount of data
[17]	ResNet-50 and GoogLeNet	ResNet results: Accuracy: 86.0, AUC 0.74; GoogleNet Results: Accuracy: 85.0, AUC: 0.75	ResNet performs better in terms accuracy and GoogleNet performs better in terms of AUC
[18]	ResNet50 UDAGAN	For DRISHTI-GSto Private dataset2: Accuracy: 89.58%; For REFUGE dataset: Accuracy: 89.75%; For the PrivateDataset1: Accuracy: 82.00%	The proposed method is memory efficient but still does not give better performance than other models [16, 19]
[19]	DeepCNN	Ensemble Classifier results: Acc.: 99.5; NasNet resultsAcc.: 99.3	NasNet performs better in terms of a single classifier. Ensemble classifier outperforms others in terms of accuracy

**Table 4** Accuracy comparison on AMD Fundus image dataset

Ref. No	Techniques used	Results	Remarks
[20]	Multiple CNN's; Random forest; Data augmentation	Overall accuracy 63.3%	The accuracy rate somehow doesn't increase
[21]	A 14-layer CNN	Accuracy: 95.45%	Better performance with less GPU utilization
[22]	DeepSeeNet	Accuracy: 67.1	Proposed method giving poor performance
[23]	VGG16	Accuracy: 90.86	The proposed method is memory efficient and gives an average performance

graph accuracy comparative analysis of FUNDUS image classification on Glaucoma and AMD respectively. Table 5 shows the comparison between all the parameters of FUNDUS image classification on the Multilabel dataset. Figure 5 represents the bar graph AUC and F1-Score comparative analysis of FUNDUS image classification on the Multilabel dataset. In Fig. 1, Das et al. show the highest accuracy of 99.6% on the Kermany dataset for the 4-class problem (CNV, DME, Drusen, Normal), and Thomas et al. shows the highest accuracy of 99.60% on the Mendeley dataset for the 3-class problem (Normal, dry AMD, wetAMD). In Fig. 2, Wan et al. show the highest accuracy of 95.68% on the kaggle2 dataset. In Figs. 3 and 4, Rehman et al. and Tan et al. show the highest accuracy of 99.5 and 95.45% on Glaucoma and AMD

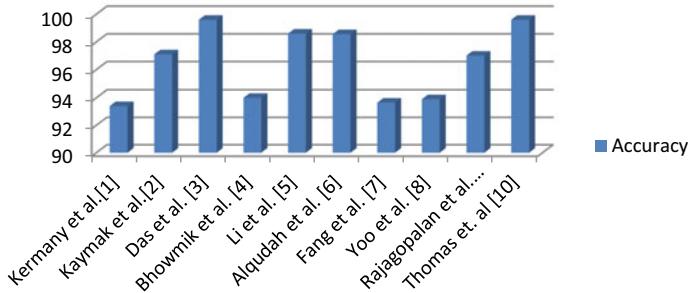
**Table 5** Accuracy comparison of multilabel fundus image dataset

Ref. No	Techniques used	Results	Remarks
[24]	CNN architecture, No fusion Techniques were used	AUC value: 80.5; F1-score: 0.85	Scarcity of a fully autonomous system to provide multi-label classification on the ODIR-19 dataset
[25]	ResNet101 with Spatial Correlation module; Image Augmentation	AUC: 0.93; F1Score: 91.3	ResNet-101 outperforms other models in terms of AUC and F1-score on the ODIR19 dataset
[26]	VGG-16 on two models, firstly on the merged and secondly on the separate left and right eye and then combined for classification	AUC: 84.93; F1Score: 85.57	However, after a given number of training epochs, the accuracy rate somehow doesn't increase on the ODIR-19 dataset
[27]	Ensemble technique such as EfficientNetB3, InceptionResNetV2	AUC: 0.74; F1Score: 0.89	Other measures such as family, gender, age, and history can be analyzed for suspected FUNDUS diseases for the diagnosis of FUNDUS diseases on the ODIR-19 dataset
[28]	Vgg, ResNet, SE-ResNet, Inceptionv4, CaffeNet and DenseNet	AUC: 0.8691; F1Score: 87.93	Inception-v4 and ResNeXt-50 possess the capability to integrate elements from various dimensions to provide a superior outcome on ODIR-19 dataset
[29]	DCNN	AUC: 0.9984; F1 score: 0.923	Better performance due to large number of images on 249,620 custom images, 39 classes

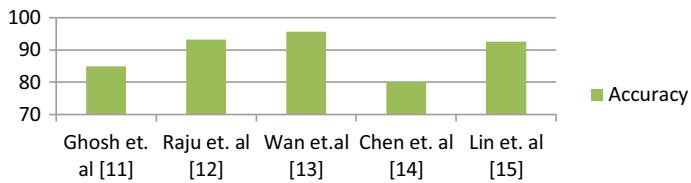
eye disease respectively. In Fig. 5, He et al. show the highest AUC of 93.0% and F1-Score of 91.3% on the multilabel ODIR-19 dataset, and Cen et al. shows the overall highest AUC of 99.84% and F1-Score of 92.3% for the multilabel classification of images.

## 4 Conclusion

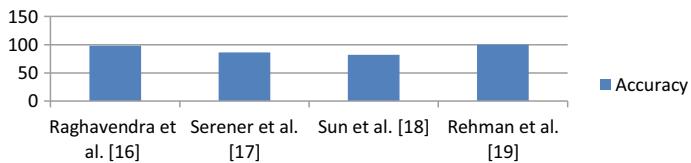
From the literature survey and comparative analysis, it is found out that DL models classify FUNDUS and OCT image classification and detection easily with greater Accuracy. This paper delivers a brief introduction to the use of medical images in the



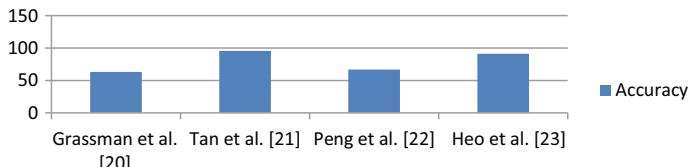
**Fig. 1** Comparison of results on Kermany and Mendeley dataset



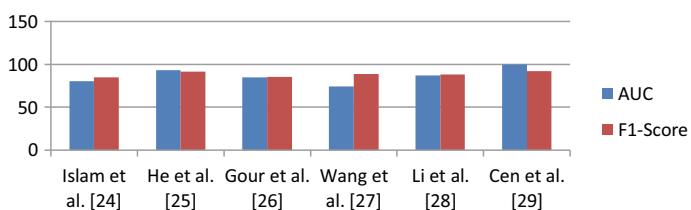
**Fig. 2** Comparison of results on multiclass Kaggle2 DR dataset



**Fig. 3** Comparison of results on Glaucoma eye disease



**Fig. 4** Comparison of results on AMD eye disease



**Fig. 5** Comparison of results on multilabel dataset

automatic classification of OCT and FUNDUS images. Different authors proposed different classification approaches depending on DL framework and the performance of the different classification methods shown by different authors were compared based on the similar dataset of the multiclass dataset with accuracy and multilabel dataset with AUC and F1-Score. Among them, Thomas et al. and Das et al. achieved higher accuracy on Mendeley and Kermany datasets respectively. Rehman et al. and Tan et al. achieved higher accuracy on Glaucoma and AMD eye disease respectively. Also, He et al. and Wan et al. achieved higher AUC and F1-score values on multilabel ODIR-19 dataset and higher accuracy percentage on multiclass kaggle2 dataset respectively. Cen et al. shows the overall highest AUC and F1-score for the multilabel classification of eye diseases. Bali et al. shows the highest validation accuracy for the multilabel classification. In future work, we will design the DCNN model which classifies FUNDUS and OCT images and perform detection.

## References

1. D.S. Kermany, M. Goldbaum, W. Cai et al., Identifying medical diagnoses and treatable diseases by image-based DL. *Cell* **172**(5), 1122–1131.e9 (2018). <https://doi.org/10.1016/j.cell.2018.02.010> (PMID: 29474911)
2. S. Kaymak, A. Serener, Automated age-related macular degeneration and diabetic Macular Edema detection on OCT Images using DL,in *2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)* (2018), pp.265–269. <https://doi.org/10.1109/iccp.2018.8516635>
3. V. Das, S. Dandapat, P.K. Bora, Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images. *Biomed. Signal Process. Control* **54**, 101605 (2019). <https://doi.org/10.1016/j.bspc.2019.101605>
4. A. Bhowmik, S. Kumar, N. Bhat, Eye disease prediction from optical coherence tomography images with transfer learning. *Commun. Comput. Inf. Sci.* **1000**, 104–114 (2019). [https://doi.org/10.1007/978-3-030-20257-6\\_9](https://doi.org/10.1007/978-3-030-20257-6_9)
5. F. Li, H. Chen, Z. Liu, X. Zhang, Z. Wu, Fully automated detection of retinal disorders by image-based DL. *Graefe's Archive Clin. Exp. Ophthalmol.* (2019). <https://doi.org/10.1007/s00417-018-04224-8>
6. A.M. Alqudah, AOCT-NET: A convolutional network automated classification of multiclass retinal diseases using spectral-domain optical coherence tomography images. *Med. Biol. Eng. Compu.* **58**(1), 41–53 (2019). <https://doi.org/10.1007/s11517-019-02066-y>
7. L. Fang, Y. Jin, L. Huang, S. Guo, G. Zhao, X. Chen, Iterative fusion convolutional neural networks for classification of optical coherence tomography images. *J. Vis. Commun. Image Represent.* (2019). <https://doi.org/10.1016/j.jvcir.2019.01.022>
8. T.K. Yoo, J.Y. Choi, H.K. Kim, Feasibility study to improve DL in OCT diagnosis of rare retinal diseases with few-shot classification. *Med. Biol. Eng. Compu.* **59**(2), 401–415 (2021). <https://doi.org/10.1007/s11517-021-02321-1>
9. N. Rajagopalan, A.N. Josephraj, E. Srithaladevi, Diagnosis of retinal disorders from optical coherence tomography images using CNN. *PloS One* **16** (7), e0254180 (2021). <https://doi.org/10.1371/journal.pone.0254180>
10. A. Thomas, P.M. Harikrishnan, A.K. Krishna, K.P. Palinsamy, V.P. Gopi, Automated detection of age-related macular degeneration from OCT images using multipath CNN. *J. Comput. Sci. Eng.* **15**(1), 34–46 (2021). <https://doi.org/10.5626/JCSE.2021.15.1.34>

11. R. Ghosh, K. Ghosh, S. Maitra, Automatic detection and classification of diabetic retinopathy stages using CNN, in *2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)* (2017). <https://doi.org/10.1109/spin.2017.8050011>
12. M. Raju, V. Pagidimarri, R. Barreto, A. Kadam, V. Kasivajjala, A. Aswath, Development of a DL algorithm for automatic diagnosis of diabetic retinopathy. *Stud. Health Technol. Inf.* **245**, 559–563 (2017)
13. S. Wan, Y. Liang, Y. Zhang, Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Comput. Electr. Eng.* **72**, 274–282 (2018). <https://doi.org/10.1016/j.compeleceng.2018.07.042>
14. H. Chen, X. Zeng, Y. Luo, W. Ye, Detection of diabetic retinopathy using deep neural network, in *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)* (2018). <https://doi.org/10.1109/icdsp.2018.8631882>
15. G.M. Lin, M.J. Chen, C.H. Yeh, Y.Y. Lin, H.Y. Kuo, M.H. Lin, M.C. Chen, S.D. Lin, Y. Gao, A. Ran, C.Y. Cheung, Transforming retinal photographs to entropy images in DL to improve automated detection for diabetic retinopathy. *Hindawi J. Ophthalmol.* (2018)
16. U. Raghavendra, H. Fujita, S.V. Bhandary, A. Gudigar, J.H. Tan, U.R. Acharya, Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf. Sci.* **441**, 41–49 (2018). <https://doi.org/10.1016/j.ins.2018.01.051>
17. A. Serener, S. Serte, *Transfer Learning for Early and Advanced Glaucoma Detection with Convolutional Neural Networks* (IEEE, 2019), pp. 271–274
18. Y. Sun, G. Yang, D. Ding, G. Cheng, J. Xu, X. Li, A GAN-based domain adaptation method for glaucoma diagnosis, in *2020 International Joint Conference on Neural Networks (IJCNN)* (2020). <https://doi.org/10.1109/ijcnn48605.2020.92073>
19. A.U. Rehman, I.A. Taj, M. Sajid, K.S. Karimov, An ensemble framework based on deep CNNs architecture for glaucoma classification using FUNDUS photography. *Math. Biosci. Eng. (MBE)* **18**(5), 5321–5346 (2021)
20. F. Grassmann, J. Mengelkamp, C. Brandl, S. Harsch, M.E. Zimmermann, B. Linkohr et al., A DL algorithm for prediction of age-related eye disease study severity scale for age related macular degeneration from color fundus photography. *Ophthalmology* **125** (9), 1410–1420. (2018). <https://doi.org/10.1016/j.ophtha.2018.02.037>
21. J.H. Tan, S.V. Bhandary, S. Sivaprasad, Y. Hagiwara, A. Bagchi, U. Raghavendra et al., Age-related macular degeneration detection using deep convolutional neural network. *Future Gener. Comput. Syst.* **87**, 127–135 (2018). <https://doi.org/10.1016/j.future.2018.05.001>
22. Y. Peng, S. Dharssi, Q. Chen, T.D. Keenan, E. Agrón, W.T. Wong, E.Y. Chew, Z. Lu, DeepSeeNet: A DL model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology* **126**(4), 565–575 (2019). <https://doi.org/10.1016/j.ophtha.2018.11.01>
23. T.-Y. Heo, K.M. Kim, H.K. Min, S.M. Gu, J.H. Kim, J. Yun, J.K. Min, Development of a deep-learning-based artificial intelligence tool for differential diagnosis between dry and neovascular age-related macular degeneration. *Diagnostics* **10**(5), 261 (2020). <https://doi.org/10.3390/diagnostics10050261>
24. M.T. Islam, S.A. Imran, A. Arefeen, M. Hasan, C. Shahnaz, Source and camera independent ophthalmic disease recognition from FUNDUS image using neural network, in *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICS CON)* (2019). <https://doi.org/10.1109/spicscon48833.2019.9065162>
25. J. He, C. Li, J. Ye, Y. Qiao, L. Gu, Multi-label ocular disease classification with a dense correlation deep neural network. *Biomed. Signal Process. Control* **63**, 102167 (2021). <https://doi.org/10.1016/j.bspc.2020.102167>
26. N. Gour, P. Khanna, Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. *Biomed. Signal Process. Control* **66**, 102329 (2021). <https://doi.org/10.1016/j.bspc.2020.102329>
27. J. Wang, L. Yang, Z. Huo, W. He, J. Luo, Multi-label classification of FUNDUS images with efficient net. *IEEE Access* **8**, 212499–212508 (2020). <https://doi.org/10.1109/access.2020.3040275>

28. N. Li, T. Li, C. Hu, K. Wang, H. Kang, *A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection*. Lecture Notes in Computer Science (including Subseries Lecture Notes Artificial Intelligence, Lecture Notes Bioinformatics), 12614 LNCS, pp. 177–193 (2021). [https://doi.org/10.1007/978-3-030-71058-3\\_11](https://doi.org/10.1007/978-3-030-71058-3_11)
29. L.P. Cen, J. Ji, J.W. Lin, S.T. Ju, H.J. Lin, T.P. Li, Y. Wang, J.F. Yang, Y.F. Liu, S. Tan, L. Tan, D. Li, Y. Wang, D. Zheng, Y. Xiong, H. Wu, J. Jiang, Z. Wu, D. Huang, T. Shi T et al., Automatic detection of 39 FUNDUS diseases and conditions in retinal photographs using deep neural networks. *Nat. Commun.* **12** (1), 4828. <https://doi.org/10.1038/s41467-021-25138-w>
30. A. Bali, V. Mansotra, Transfer learning-based one versus rest classifier for multiclass MultiLabel ophthalmological disease prediction. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, **12** (12), 537 546 (2021). <https://doi.org/10.14569/IJACSA.2021.0121269>

# A Secure Framework for the Deployment of Microservices Using Cloud Container Technology



**Salini Suresh, N. Ramachandran, M. Hanumanthappa, K. Ravikumar, and Anubha Jain**

**Abstract** Nowadays, companies have been focusing on application transformation due to many issues in the existing monolithic application. Some of the notable issues are that while upgrading a single module, the entire application has to be down. During peak business hours, the software crashes or does not respond due to heavy load. To overcome these issues, companies are primarily focusing on microservices using containers with adaptability across the framework. Specific microservices can still be replicated on demand or simplified to fit the needs, even as requirement ranges. The ultimate infrastructure must be optimized to manage the complexity of individual microservices and make an adequate range of available instances to optimize efficiency to accommodate such resiliency. Although container technology is widely used in improving microservices, the security vulnerabilities in microservice are high compared with a monolithic application. This article describes the security vulnerabilities of containers for use in microservices and proposes a secure framework that enhances the isolation of containers.

**Keywords** Cloud · Microservice · Security · Container technology · Monolithic

---

S. Suresh (✉)

Department of Computer Applications, Dayananda Sagar College, Bangalore, India  
e-mail: [pnsalinisuresh@gmail.com](mailto:pnsalinisuresh@gmail.com)

N. Ramachandran

Indian Institute of Management, Kozhikode, India

M. Hanumanthappa

Department of Computer Science and Applications, Bangalore University, Bangalore, India

K. Ravikumar

Kalinga Institute of Industrial Management, Bhubaneshwar, India

A. Jain

IIS (Deemed to be University), Jaipur, India

## 1 Introduction

The way an application is built has changed over the years. Earlier a software was built or developed from scratch by a team or set of coders writing a shared code. Instead of endless codes written by developers for various applications, the already written codes are used as a base for developing new applications as required by the developers, which helps in saving lots of time. The development platform on the cloud provides libraries or an ecosystem of codes that the developers can easily access and integrate into an application. In the present scenario, developers distribute themselves to separate teams that will work on particular parts or a section of the application instead of building an entire application by one team. These smaller amounts of codes written for a particular part of an application are known as microservices.

In the monolithic application, the modules are tightly coupled, so they cannot be executed separately. In microservices, the functional system can be decomposed into manageable and independently deployable components with enhanced agility for business as the desired outcome [1, 2]. Microservices enable developers to create applications from various individual components that can be effectively modified, removed, or redesigned without affecting the whole application, which is impossible in monolithic applications. Monolithic applications which are transformed into service oriented microservice architectures radically change the perceptions of application development.

### 1.1 *Overview of Microservices*

Microservices break large applications into their core function, such as search recommendations, customer reviews, product catalogues, etc. It is developed independently yet can work together on a cloud development platform to create a functioning platform. These microservices are distributed to the particular or required systems by containers. So if one microservice does not work for an application, that particular service is discarded by the developer and develops a new or a compatible microservice. Since the developer can replace only a particular part of an application, it doesn't affect the rest. For these microservices to work together, they use service discovery to find each other, which creates a road map for many microservices to communicate. It communicates by combining Message Brokers, Event Streaming, and Application Programming Interfaces (API). Since the developers use microservices instead of creating code from scratch, it becomes easy to add new functionality.

This architecture is an approach in which one application is composed of many single deployable services or components which have a stack running on their containers. For instance, this means that the components of an application can be developed and maintained more efficiently by multiple developers working on the different components. These developers work independently and can use different runtime environments and stacks for the individual components. Components facing

too much load can be scaled separately or individually rather than scaling the entire application, which would reduce the effective cost. Although one of the significant motivations for developing a microservice architecture is to solve scalability, a difficult challenge for microservices is preserving the security while providing the required scalability dynamically.

## ***1.2 Overview of Container***

Compared to conventional hypervisor virtualization, the container technique uses lightweight virtualization and has many performance upgrades. Unlike a Virtual Machine (VM) that virtualizes the entire OS, the features in the host OS kernel are used to restrict cycle-level applications for both applications. The essential benefit of this strategy is the avoidance of duplication of instructions and a substantial improvement in performance since instructions are executed on the host. The overhead generated by the hypervisor is eliminated by using containers over hypervisors, with effective isolation [3, 4]. Due to this, containers are an attractive choice for Platform as a Service (PaaS), substantially reducing costs by decreasing virtualization overhead [5]. A framework can effectively deploy a more significant number of containers than the number of virtual machines deployed. It is also cost-effective to deploy frameworks with many containers than many frameworks with scarcely any VMs. Moreover, removing the need to boot an entire OS increases the start-up time significantly [6].

## ***1.3 Container for Microservices***

Containers being lightweight and portable, it provides a microservice-friendly environment. Each container can be created with the application and dependencies required to run [7, 8]. Subsequently, the container runs on the host operating system as an isolated process, with the kernel being shared with the other containers. Therefore, the container is placed in a VM environment. In addition, resource utilization benefits from using virtual machines; containerization makes it further portable and efficient. Kubernetes [9] made the executives in massive clusters more agreeable by actualizing a steady REST API. Kubernetes is an extensible, open-source platform for managing containers. Kubernetes enables the containers to run in massive clusters by actualizing a steady REST API.

## 2 Problem Statement

The application inside the containers can become malicious or semi-honest. Protecting the containers from each other in a cluster is a significant challenge. In high risk attacks like meltdown [10], attackers with permissions can run malicious commands. In this attack, the adversary leaks the kernel information from the host OS and all other hosted containers. The malicious container can access confidential data of other containers and learn resource utilization, thus affecting the integrity of the application. Mitre [11] is another threat to containers, which causes applications to access random locations in their memory. Due to a lack of inter container security, malicious containers hosting the microservice can lead to Denial of Service (DoS). The compromised images in the registry may also lead to resource hijacking.

## 3 Literature Review

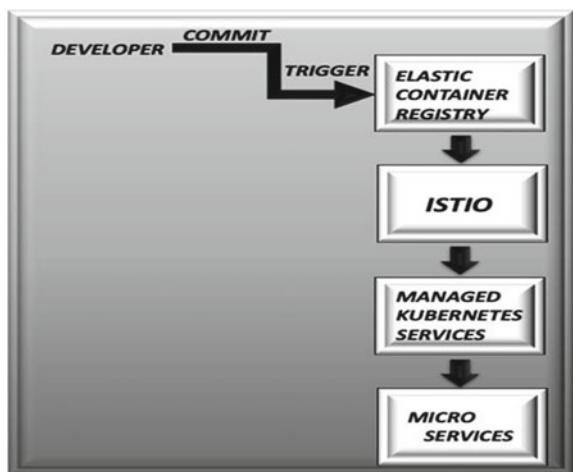
Chelladurai et al. [12] have discussed the security concerns of Docker containers and presented a detailed review of the existing works on the same. Authors have listed down security approaches prevalent today to mitigate the security vulnerabilities of Docker. The authors have presented a theoretical discussion of a proposed method to encounter DoS and conducted the experimental study by executing the proposed model for various test cases. Thangapandiyan et al. [13] have proposed a framework that provides privacy, trust and security from adversaries. The paper's focus is delivering a trust management model as a service for the providers and the users. The authors noted that the suitability of the proposed framework was investigated with real-time user responses. Pasquier et al. [14] have discussed the isolation provided by container technology and hypervisor-based virtual machines. The proposed IFC (Information Flow Control) empowered cloud model offers secure and elastic data sharing, and the paper has carried out a comparison of isolation.

## 4 Methodology

The proposed approach provides a secured framework to run microservices in a container by creating a DevOps environment. The proposed microservices architecture addresses the security challenges arising from a remarkable upsurge in the number of components deployed and interacting with hybrid network environments.

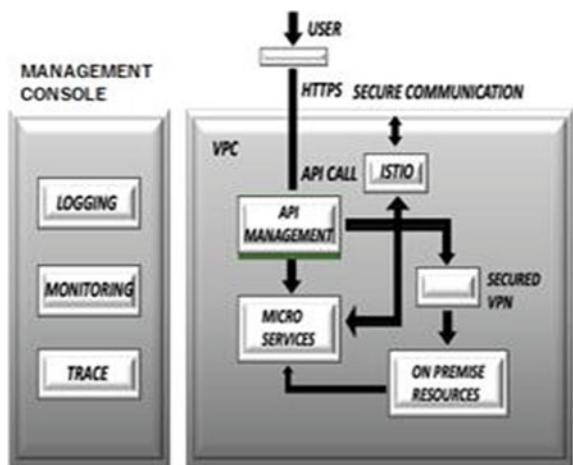
In this approach, when the developer creates a new instance and triggers it, the container registry will update the concerned profile, security and access control. The service mesh or ISTIO manages the resource and container allocation, and container level security is applied to the user. Kubernetes manage the container orchestration. Figure 1 illustrates the DevOps environment to run microservice.

**Fig. 1** DevOps environment for microservice



The API call triggers the service mesh when the user accesses the Virtual Private Cloud (VPC) through a secured VPN. Figure 2 illustrates the overview of microservices. Even though APIs are gateways that provide access to microservices, endpoint APIs are vulnerable to attack and lead to unauthorized access to sensitive data. The management console strictly monitors and traces the interactions, and network access control lists are limited to the IP addresses required to manage the cluster. The Kubernetes API deployed in this framework regulates the access by RBAC (Role-based access control).

**Fig. 2** Overview of microservice architecture



**Table 1** Number of users and response time of the monolithic application hosted on-premise with high availability

Monolithic application hosted on-premise	
No. of users	Login response-time (milliseconds)
10	15
100	654
500	1876
1000	3783
5000	7345
10,000	10,487

#### 4.1 Overview of an Existing Setup

An e-commerce application runs in a retail boutique, and the application was developed by using .Net as a front-end tool and PostgreSQL as a back-end database. This application runs on a server with 4 Intel Xeon processors with 24 cores and 172 GB RAM. One more similar configuration server runs for high availability, both working in a clustered environment. The problem with the legacy application is during peak hours number of customers is more. During that time, the application's performance goes down, leading to either no response from the application or system hangs, especially while logging in. Even after increasing the hardware resources, there was no improvement in performance. Table 1 shows the number of users and response time of the monolithic application hosted on-premise with high availability.

#### 4.2 Experiment

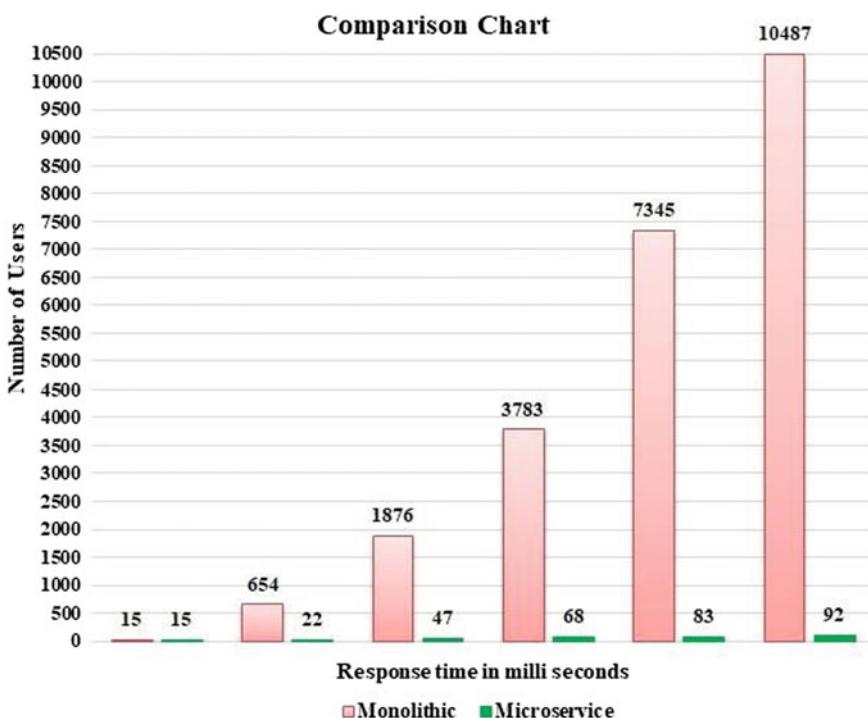
The existing monolithic application has many modules such as login, purchase, stock, billing, etc. In these modules, the user faced login issues during peak hours even after fine-tuning the application and upgrading the hardware resources. Therefore in this experiment, the login module was rewritten in PHP as a microservice using container technology, and it is hosted on the cloud. It is managed through the google Kubernetes engine, and this microservice-enabled e-commerce application has been set with a scale of upto 20 pods. It runs on google public cloud with 24 virtual CPUs, 326 GB RAM, and without high availability. This experiment tested with various workloads, various numbers of users, and each user's access time was 3 min of the application. Table 2 shows the number of users and response time of the microservice enabled application hosted on the cloud without high availability.

**Table 2** Number of users and response time of the microservice-enabled application hosted on the cloud without high availability

Microservice application hosted on cloud	
No. of users	Login response time (milliseconds)
10	15
100	22
500	47
1000	68
5000	83
10,000	92

### 4.3 Data Analysis

Figure 3 illustrates the comparison chart of both monolithic applications and microservice applications response time. It clearly shows that the monolithic applications hosted on-premise with high availability and Microservice application hosted on the cloud without high availability, the response time is much lesser in microservice applications.



**Fig. 3** Comparison of both monolithic applications and microservice applications response time

## 5 Conclusion and Future Scope

Microservice is a proven technology only hitch in this technology is security. The proposed methodology secure framework is included in the microservice application and tested in a real-time environment. Compared with the monolithic applications response time with the proposed secured microservice application, it has shown a good response time, which is more significant.

As per the analysis, if the user is less, there is no difference in the response time. When the number of users increases, the response time in the proposed framework is significantly less. If the number of users is very high, the response time with the monolithic software and the proposed framework is enormous. It could be due to the integration of three technologies: cloud, microservice, container, and implemented only a single module.

This study has considered only a single module and without high availability. Therefore, it is suggested that the entire monolithic application be split into microservices and implement the same experiment on the cloud with high availability and test it out.

## References

1. N. Dragoni et al., Microservices: Yesterday, today, and tomorrow, in *Present and Ulterior Software Engineering* (Springer, Cham, Switzerland, 2017), pp. 195–216
2. B. Golden, *3 Reasons Why You Should Always Run Microservices Apps in Containers*. Accessed: 11 Nov 11 2017 (2016)
3. S. Suresh, L.M. Rao, A container based collaborative research cloud framework with a hybrid access control approach for enhanced isolation, in *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, Mysuru (2017), pp. 1–6. <https://doi.org/10.1109/ICEECCOT.2017.8284605>
4. S. Suresh, L. Rao, CCCORE: Cloud container for collaborative research. *Int. J. Electr. Comput. Eng.* **8**, 1659–1670 (2018)
5. R. Dua, A.R. Raja, D. Kakadia, Virtualization vs. containerization to support paas, in *2014 IEEE International Conference on Cloud Engineering (IC2E)* (IEEE, 2014), pp. 610–614
6. B. Xavier, T. Ferreto, L. Jersak, Time provisioning evaluation of KVM, Docker and Unikernels in a cloud platform, in *Proceedings—2016 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2016* (2016), pp. 277–280
7. S. Vaucher, R. Pires, P. Felber, M. Pasin, V. Schiavoni, C. Fetzer, SGX-Aware container orchestration for heterogeneous clusters, in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)* (July 2018), pp. 730–741
8. Secured containerization for enhanced security and isolation in collaborative research cloud. *Int. J. Appl. Eng. Res.* ISSN 0973-4562 (vol. 14, no. 15) (2019), UGC Approved Journal—2017 (Journal No.—64529)
9. The Linux Foundation R. *Kubernetes—Production-Grade Container Orchestration*. <http://kubernetes.io/> (2016)
10. M. Lipp et al., Meltdown: Reading Kernel memory from user space, in *27th USENIX Security Symposium (USENIX Security 18)* (2018), pp. 973–990
11. P. Kocher et al., *Spectre Attacks: Exploiting Speculative Execution* [Online]. Available: <https://arxiv.org/abs/1801.01203> (2018)

12. J. Chelladhurai, P. Chelliah, S. Kumar, Securing Docker containers from denial of service (DoS) Attacks. *IEEE Int. Conf. Serv. Comput. (SCC)* **2016**, 856–859 (2016)
13. M. Thangapandian, P.M. Rubesh Anand, A secure and reputation based recommendation framework for cloud services, in *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (2016), pp. 1–4
14. J. Bacon, D. Eyers, T.F.J.-M. Pasquier, J. Singh, I. Papagiannis, P. Pietzuch, Information flow control for secure cloud computing, in *IEEE Transactions on Network and Service Management*, vol. 11, issue 1 (2014), pp. 76–89

# Time Series Analytics—A Fuel to Data Science in Forecasting Analytics for Time-Stamped Data



Sunitha, Salini Suresh, Anubha Jain, and V. Suneeta

**Abstract** Building an accurate forecasting model on time-dependent data for various Business solutions like Occupancy rate, Sales Forecasting, Stocks Prediction, Pairs Trading, and other various portfolio management problems is always an uphill task. Time series analysis is one of the frequently preferred ML algorithms for addressing such challenges. Many statistical and neural network-based models like RNN have been proposed for analyzing and forecasting time series datasets. This paper discusses the importance of data pre-processing for time-stamped data by understanding the trend, seasonality, and random factors. Also tried and tested Augmented Dickey-Fuller (ADF) test to identify the presence of stationarity in the data. We also discussed Time Series Techniques like ARIMA, HOLTWINTERS. We developed a forecasting model to understand business objectives to split the data and evaluate the model. We evaluated the results between both models and tested the models with Information criteria techniques for better results.

**Keywords** Time series forecasting · ARIMA · Stationarity · Seasonality · AIC test

## 1 Introduction

Data Science incorporates computational, analytical, and statistical skills to extract useful information from data. It applies Artificial Intelligence techniques and Machine Learning Algorithms to numeric, textual, or multimedia data. Data Science combines a range of skills to analyze different types of data collected from the Internet, Sensor devices, Smart Phones, and various other sources that derive actionable insights. The collected data encompasses Analytics which consists of cleaning, grouping, and influencing the data to execute complex data analysis. Data Science

---

Sunitha (✉) · S. Suresh · V. Suneeta  
Dayananda Sagar College of Arts, Science and Commerce, Bangalore, India  
e-mail: [sunitha6329@gmail.com](mailto:sunitha6329@gmail.com)

A. Jain  
IIS (Deemed to be University), Jaipur, India

also incorporates areas like data engineering, data mining, predictive analytics, data visualization, etc.

### ***1.1 Importance of Data Science in Prediction***

Data is essential for any organization since it permits them to make decisions based on facts, trends, and statistics. As the impact of data, data science has advanced into a multi-disciplinary business. Predictive analytics is used to assist organizations in forecasting future outcomes based on historical data. It's a type of data analytics that focuses on forecasting future consequences [1].

### ***1.2 Forecasting and Prediction***

Forecasting and prediction are both concerned with the future. It is the way of evaluating a future state in respect to any type of operation. The approach can use current or historical data to forecast future events [2].

Forecasting is the process of making predictions within different instances at different levels, is carried out by senior authorities, analysts, and skilled statisticians. This is done both qualitatively and quantitatively. Quantitative forecasting is an analytical approach to making forecasts that aim to compare variables using previous records and trends. Quantitative Forecasting is used in Time Series Analysis. To make forecasts, qualitative forecasting depends on judgment instead of numerical metrics.

A Prediction is a testimony that aims to justify a potential outcome or opportunity event. Public and private organizations utilize predictive analytics to determine the forecasting needs to fulfill the futuristic growth in numbers. Forecasting is a time-based prediction appropriate to deal with time-series data. Whereas the prediction does not depend on time but can be based on various factors that deal with target variables.

### ***1.3 Forecasting Techniques***

ARIMA model is generally used for forecasting in such models that considers non-zero autocorrelation [3]. The operation of ARIMA can be described in three steps [4]. In the beginning, the Identification stage is the first step in which time-series data directs towards stationarity using ACF (Auto Correlation) and PACF (Partial autocorrelation).

The next step is the estimation stage which lets us assess the model's sustainability. With the help of the estimate stage, the next phase which is the forecasting stage may

identify the best forecasting output by way of future time series values as well as upper and lower bounds for forecasts. The Holt-Winters Exponential smoothing is used for forecasting time-series data with both a trend and seasonal variance. The Holt-Winters approach is made up of the four forecasting methodologies which are Weighted Averages, Exponential Smoothing, Holt Exponential, and Holt-Winters Exponential Smoothing.

## 2 Related Work

In [4] authors discuss how time series forecasting is critical for constructing an effective forecasting model for predicting the spread of this infectious disease in the future. The ARIMA and Holt-Winters models are utilized in this work to construct a 20-day forecast model to foresee the effect of the COVID-19 Pandemic. The Automatic ARIMA models for predicting the returns of stock price moment-by-moment data are discussed in [5].

In [6] Authors have proposed a customized (ARIMA) integrated modeling technique to depict the correlation and probable allocation of determined time-collection records. In [7] analysis produces a fusion architecture for predicting the numbers for electricity costs. The incentive of the simple-to-apply and relatively easy-to-tune ARIMA model is mixed with the estimated power of district intuitive methods in the suggested architecture. The architecture is more dependable and precise than the particular forecasting approaches. The [8] paper focuses on the use of Holt-Winters exponential smoothing algorithms to evaluate seasonal time-series data.

## 3 Methodology: Forecasting Air Passengers Occupancy with Time Series Modelling Using ARIMA and HOLT WINTERS

The objective is to build a model to forecast future occupancy figures. Forecasting models will enable companies to make themselves prepared for the future and to develop Business plans accordingly.

This data is all about anonymous airline occupancy information during the years 1949 and 1960 and it's open-source data available in R. It contains the data for every month starting from Jan to Dec for the mentioned years.

### **3.1 Data Interpretation and Preprocessing**

In Time Series Analysis Stationarity is the main characteristic. It refers to the process of time series generating which is constant. It does not depend on the observed time and at distinct times the trend and seasonality will affect the covariance of the time series. Like the mean, variance, and autocorrelation structure are constant over time Stationarity plays a very important role. The tools used in forecasting are constant. Ex. ARIMA. When forecasting the future, most time series models are developed in an assumption that every point is independent of one another. The common assumption of all the techniques assumes the data is stationary. The Non-Stationary refers to the data in which the Mean, Variance, and Correlation change over time because of the factors like Trend, Seasonality, Noise, and random Walk.

- plot (AirPassengers),
- cycle (AirPassengers),

- adf. test (AirPassengers), pp. Test (AirPassengers).

## **4 Experiment**

With respect to the given Air Passengers example, the above code is used to identify the data is stationary or not using Adf and PP Test.

**Augmented Dickey-Fuller Test:** This test belongs to a specific category test which is also called the “Unit Root Test”. The Dicker-Fuller test is used for testing the null hypothesis of the presence of unit root. The ADF tests include high order Regressive process in the model which adds more thoroughness to the test.

**Phillips-Perron Test (PP Test):** In a Univariate Time series, the PP test considers the null hypothesis of a Unit Root and performs Least- Squares Regression (OLS/WLS) to estimate coefficients in the Model. The tests follow nonstandard distributions under null asymptotically.

### **4.1 Data Decomposition: Steps Required in Decomposition**

I. Calculate the Trend. Mainly there were two areas

(I) Calculate the Trend. Mainly there were two areas. (a) Smoothing Procedure and (b) Identify the equation of regression for the trend. (II) The series De-Trending: There were two main models used to decompose the time series. Additive Model and Multiplicative Model. The Additive decomposition subtracts the trend estimates whereas Multiplicative decomposition divides the series by the trend estimated values. (III) From the De-Trending series, the seasonal values can be calculated. Firstly, estimate the Mean for a specific period and later the seasonality effects could be estimated along with the trend by the regression equation. That is the number of

seasonal factors is equal to the frequency of the series. (IV) The Seasonality should normalize by any one of the decomposition models. From the Additive model, the effects are adjusted in such a way that the average of ‘d’ components is 0. From the Multiplicative model, the effects are adjusted so that the average ‘d’ is equal to 1. (V) The Residuals should be estimated. That is finding the irregular component. (VI) After decomposing the time series, a specific Model is chosen to fit exactly for stationary Residuals.

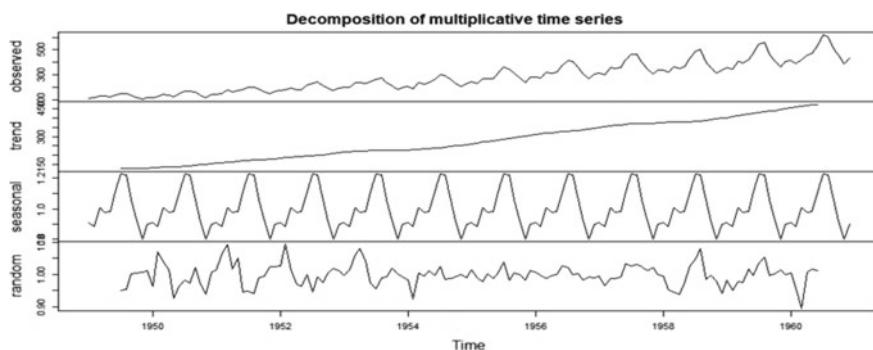
## 4.2 Modelling

In this paper, we build a Model through ARIMA and HOLTWINTER Model [9]. While testing for the stationarity we performed decomposition by considering the multiplicative component instead of the additive component. This will derive the connection between factors like a trend, random walk, and seasonality function in the data whether they are additive or multiplicative.

The Air passenger’s data has been observed with the factors of a random walk, a high seasonality, and an upwards trend in the data. Figure 1 indicates all these factors in a plot.

- decompose = decompose (Air Passengers, ‘multiplicative’)
- decompose\_AP
- plot(decompose\_AP) (Table 1).

For Airpassenger data which we are modeling haven’t found any stationarity, hence will consider the undecomposed data.



**Fig. 1** The plot which represents decomposition of multiplicative time series

**Table 1** Represents the code how to decompose the non-stationarity factors of the time series data

Data without Seasonality	<code>AP_no_seasonal = AirPassengers/decompose_AP\$seasonal</code>
Data without Random Walk	<code>AP_no_random = AirPassengers/decompose_AP\$random</code>
Data without Trend	<code>AP_no_trend = AirPassengers/decompose_AP\$trend</code>
Data with Stationary	<code>AP_Stationary = AirPassengers/(decompose_AP\$seasonal decompose_AP\$random + decompose_AP\$trend)</code>
Plot the Stationary Data	<code>plot (AP_Stationary)</code>

### 4.3 Data Splitting

Split the data into 2 windows such that the train data will be from (Jan 1949 to Dec 1958) and the rest is the testing data (from Jan 1959 to Dec 1960).

- `AP1958 = ts (AirPassengers, frequency = 12, start = c (1949,1), end = c (1958,12))`.

### 4.4 Building the Model

Now the train and the test data are ready for modeling. Here will consider the train data to learn by applying the ARIMA model. This algorithm requires the 3 most important parameters p, q, and d which represent the integration of Auto Regression (AI) and Moving Average (MA). To derive those parameters automatically by the system rather than calculating manually will use [10]

`AUTO.ARIMA [11]`

- `AAModel = auto. arima (AP1958),`
- `AAModel.`

`AUTO.ARIMA` function derived the below values for Auto Regression (AR) derivative “p” using PACF and respectively for Moving Average (MA) using ACF “q” and even for diff “d”. We used Akaike Information Criterion (AIC) for selecting the p, q, and d values [12]. Even we can use the Bayesian Information Criterion (BIC). The optimal values are given by `AUTO.ARIMA` is as follows:

For order  $(p, q, d) = (1, 1, 0)$  and for seasonal order  $(p, q, d) = (0, 1, 0)$ .

Considering those values developed the Arima model by providing the train data and taking the log values into account to bring the scalability in the data.

- `AAModelx = arima (log (AP1958), order = c (1, 1, 0), seasonal = list (order = c (0,1,0), period = 12))`

Forecasting next two years (1959 and 1960) using Arima model (Table 2).

**Table 2** Represents the code for prediction of next two years ahead from training time stamp ie., Jan 1959 to Dec 1960

Nxt2AApred = pred(AAModelx, n.head = 24)	Next 2 years of Predicted data by ARIMA
Nxt2AApred = 2.718*Nxt2AApred\$pred	Converting from log to natural numbers
Nxt2AApred = round(Nxt2AApred)	Rounding off the number without decimals
Nxt2AApred	Prints the values

**HOLTWINTERS:** Forecasting demands in the models which account for trend and seasonality. The decision depends on the result of the model used. Holt Winter's is one of the popular time series prediction methods.

- HWModel = HoltWinters (AP1958)
- Nxt2HWpred = predict (HWModel, seasonal = “multiplicative”, n. ahead = 24)  
» Nxt2HWpred = round (Nxt2HWpred, digits = 0).

Developed another time series model using HOLT WINTERS.

## 5 Results and Discussion

Now comparing the actual data of 1959 and 1960 with the forecasted values of ARIMA and HOLTWINTERS models for evaluation.

According to Figs. 2 and 3, Observations show Holtwinters is showing better results when compared to ARIMA (Fig. 4).

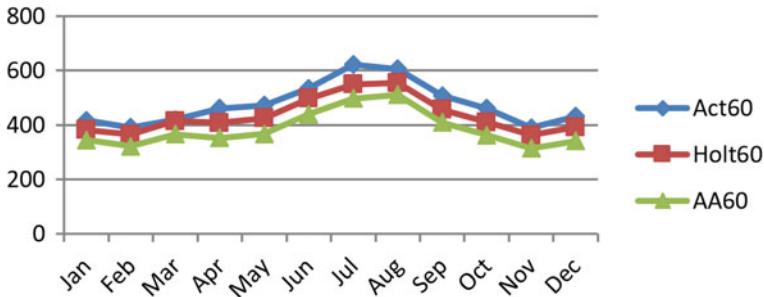
### Various KPIs for evaluation of forecasting model

- (1) **Mean Absolute Error (MAE):** MAE is a metric that measures the absolute difference between actual and predicted values.

**Merits of MAE:** The Predicted variable unit will be similar to what we get in MAE and Non-sensitivity to Outliers.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<b>Act59</b>	360	342	406	396	420	472	548	559	463	407	362	405
<b>Holt59</b>	354	337	388	380	398	470	521	528	429	383	335	365
<b>AA59</b>	342	320	364	350	365	437	494	508	406	361	312	339
<b>Act60</b>	417	391	419	461	472	535	622	606	508	461	390	432
<b>Holt60</b>	382	365	415	408	425	497	549	555	457	410	363	392
<b>AA60</b>	344	322	366	352	367	440	497	511	409	363	314	341

**Fig. 2** Shows the actual values of 1959, 1960 along with the predicted values for same years by ARIMA and HOLT WINTERS



**Fig. 3** Plot represents the comparison of the actual data and predicted values

accuracy(Nxt2Hwpred, AP5960)							
	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
Test Set	31.1667	35.7771	31.1667	6.67933	6.679327	0.472726	0.6902795
accuracy(Nxt2AApred, AP5960)							
	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
Test Set	67.9167	73.4898	67.91667	14.7968	14.79683	0.713556	1.451876

**Fig. 4** Indicates the accuracy values of both ARIMA and HOLT WINTERS results

**Demerits of MAE:** Optimization techniques like Gradient descent and Stochastic gradient descent are required because the MAE graph is not differentiable.

- (2) **Mean Squared Error (MSE):** MSE gives us a square difference between actual values and predicted values.

**The merit of MSE:** As the graph of MSE is highly differentiable, one can easily use it as a loss function.

**Demerit of MSE:** Sensitive to Outliers because it penalizes the outliers.

- (3) **Root Mean Squared Error (RMSE):** Name itself indicates, it's a simple square root of mean squared error.

**The merit of RMSE:** Loss function will be easily interpreted.

**Demerit of RMSE:** Sensitive to Outliers.

All kinds of errors like Mean Error (ME), Root Mean Square Error (RMSE), Mean Absolute Error (MSE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), and most important AIC (Akaike Information Criteria) scores are less for Holtwinters [13].

## 6 Conclusion

Time series forecasting model on AirPassengers data from Jan 1949 to Dec 1960. Tested stationarity issues in the data and found no stationarity using ADF and PP tests. Also mentioned is how to decompose the data in case there is any stationarity. Developed both ARIMA and HOLT WINTERS models by splitting the data into train and test. Observed HOLT WINTERS gives better results when compared with ARIMA by evaluating models under various accuracy parameters like MAE, MAPE, MSE, AIC, etc.

## References

1. C.V. Apte, S.J. Hong, R. Natarajan, E.P.D. Pednault, F.A. Tipu, S. Weiss, Data-intensive analytics for predictive modeling. *IBM J. Res. Develop.* **47**(1), 17–23 (2003)
2. V. Dhar, Data Science and Prediction. *Commun. ACM.* **56** (2012). <https://doi.org/10.2139/ssrn.2086734>
3. M. Kumar, M. Anand, An application of time series Arima forecasting model for predicting sugarcane production in India. *Stud. Busi. Econ.* **9**, 81–94 (2014)
4. M. Panda, Application of ARIMA and Holt-Winters forecasting model to predict the spreading of COVID-19 for India and its states. *medRxiv* n. pag (2020)
5. L. Yermal, P. Balasubramanian, Application of auto ARIMA model for forecasting returns on minute wise amalgamated data in NSE. *IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC) 2017*, 1–5 (2017). <https://doi.org/10.1109/ICCIC.2017.8524232>
6. W. Jacobs, A.M. Souza, R.R. Zanini, Combination of Box-Jenkins and MLP/RNA models for forecasting combining forecasting of box-Jenkins. *IEEE Latin America Trans.* **14** (4) (2016)
7. S.I. Vagropoulos, G.I. Chouliaras, E.G. Kardakos, C.K. Simoglou, A.G. Bakirtzis, Comparison of SARIMAX, SARIMA, Modified SARIMA and ANN-based models for short-term PV generation forecasting, in *IEEE International Energy Conference (ENERGY)*, Leuven (2016), pp. 1–6
8. P.S. Kalekar, Time series forecasting using Holt-Winters exponential smoothing (2004)
9. E. Hannan, The estimation of the order of ARIMA process. *Ann. Stat.* **8**, 1071–1081 (1980)
10. P. Meisen, M. Recchioni, D. Schulberg, T. Meisen, Modeling and processing of time interval data for data-driven decision support, in *IEEE International Conference on Systems, Man, and Cybernetics*. Dipl.-Inform., 5–8 Oct 2014
11. A. Coghlan, *A Little Book of R for Time Series*, Readthedocs.org (2010). Available online at <http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/src/timeseries.html>
12. Manoj Kumar, M. Anand, An application of time series ARIMA forecasting model for predicting sugarcane production in India. Published on 2014, *Studies in Business and Economics*, Corpus ID: 150562933
13. S. Makridakis, Time-series analysis and forecasting: an update and evaluation. *Indian Stat. Rev.* **46**, 255–278 (1978)

# Sparse Autoencoder and Deep Learning Based Framework for Multi-label Classification of Chronic Diseases



Atul Kumar Ramotra, Amit Mahajan, and Vibhakar Mansotra

**Abstract** The accurate prediction of chronic diseases at an early stage can help the patients to plan the future course of treatments effectively so as to avoid any life threatening situations. Machine learning techniques provide cost and computationally inexpensive solutions for multi-label as well as multi class disease classification problems. In case of multi-label classification problems more intelligent techniques of data analytics are required due to the presence of correlations and label dependencies in the dataset. The accuracy of prediction also depends upon the representation of the data. In this paper a Sparse Autoencoder and Deep learning based Framework has been proposed to predict the presence of multiple diseases in the patients. The framework consists of three phases. In the first phase, Sparse Autoencoder is used to learn features from the original dataset. In the second phase, Label Powerset method is used to transform the problem from multi-label to multiclass classification problem. In third and the final stage, a deep learning architecture is applied. 80.85% of accuracy has been achieved by the proposed framework.

**Keywords** Machine learning · Multi-label classification · Sparse autoencoder · Deep learning · Neural network

## 1 Introduction

Large healthcare datasets of medical field can be used to obtain important information related to the early diagnosis of various chronic diseases. An early stage prediction of chronic diseases can help the patients in planning the future course of treatments effectively to avoid any life threatening situation. Practice of medical diagnosis is a complicated process and presence of more than one disease in a patient at the same time can bring additional challenges. Various factors responsible for causing multiple diseases are often complex to understand. More intelligent techniques of data

---

A. K. Ramotra (✉) · A. Mahajan · V. Mansotra

Department of Computer Science and IT, University of Jammu, Jammu, Jammu and Kashmir 180006, India

e-mail: [ramotraatul@gmail.com](mailto:ramotraatul@gmail.com)

analytics are required to understand the complex relationships between these factors. Healthcare data related to the patients with symptoms of multiple diseases present at the same time are categorized under multi-label classification problems. The presence of correlations among the attributes and label dependencies in the dataset makes these problems more challenging than multi class problems. Various techniques offered by machine learning helps in providing cost effective and computationally inexpensive solutions for the multi-label classification problems [1].

Most of the classification algorithms cannot be applied directly on the multi-label datasets. Problem transformation methods such as Binary Relevance (BR) and Label Powerset (LP) are used to solve the multi-label classification problems which convert the multi-label classification problem to multiclass classification problem. Presence of non-linear relationships between various factors and label dependencies complicates the process of extracting meaningful interpretations from multi-label datasets. LP method is preferred over BR because of its ability to handle label dependencies [2]. The form of data to be presented to the learning systems also plays an important role in multi-label classification problems. Feature learning methods offered by machine learning can be utilized to find the data representation which can further enhance the performance of the algorithms. In recent years, Autoencoder (AE) has been used as a feature learning method by many researchers. AE learns the feature representation from the original data using unsupervised learning technique [3]. In this paper a framework has been proposed using two different neural networks for multi-label disease prediction. Sparse autoencoder (SAE) is used to represent features from the original dataset, LP method is used to transform multi-label to multiclass problem and deep neural network based framework has been used for the final classification task.

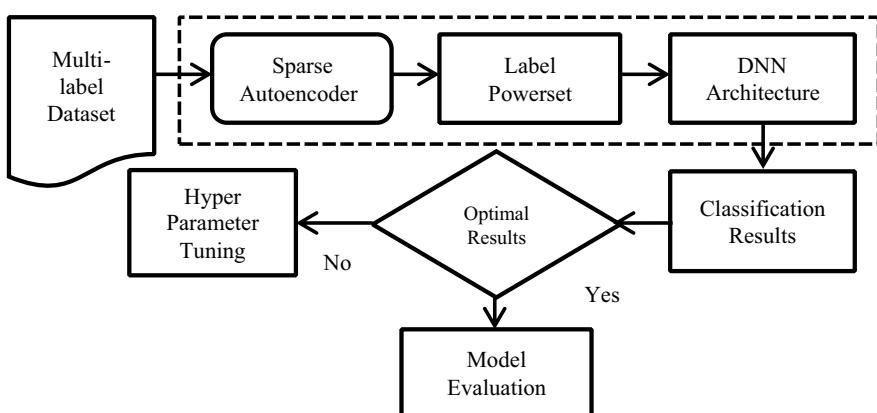
## 2 Sparse Autoencoder

AE is a neural network using feed-forward and unsupervised method to encode the input data. The basic structure of the AE contains encoding layer which encodes the input into codes using the original data and the decoding layer which decodes the input into the original form using the codes. The design of the encoding as well as the decoding layer is always symmetric that is the number of nodes in both the layers are always same. However the number of neurons in the codes may vary. If the number of neurons in codes is less than the encoding layer, it is called Undercomplete AE and if number of neurons is more than the encoding layer, it is called Overcomplete AE [4]. AE can suffer with the problem of copying the input data through the network. This problem is generally avoided using a restriction or sparsity. Sparsity is a small penalty added to the layers of an AE to avoid overfitting and copying of the original data. The AE with sparsity in the layers is known as SAE. Many researchers have used SAE to represent the features of the original dataset to increase the accuracy of the classification algorithms. Mienye et al. [5] used a two staged approach to predict the presence of heart disease. The first stage involves the representation of the

original data using SAE and the second stage includes the employment of an artificial neural network as a classifier for the disease prediction. The authors compared the results obtained from their experimental setup with the work of other researchers and claimed to have achieved the accuracy of 90%. Preethi et al. [6] designed an intrusion detection system using SAE for feature learning and dimensionality reduction of the dataset. The authors used Support Vector Regression as the final predictor in the system. Kadam et al. [7] proposed a method for the prediction of breast cancer using stacked SAE for feature ensemble learning and softmax regression for classification. The model attained 98.60% of accuracy on dataset containing 569 instances and 30 input features.

### 3 Proposed Framework

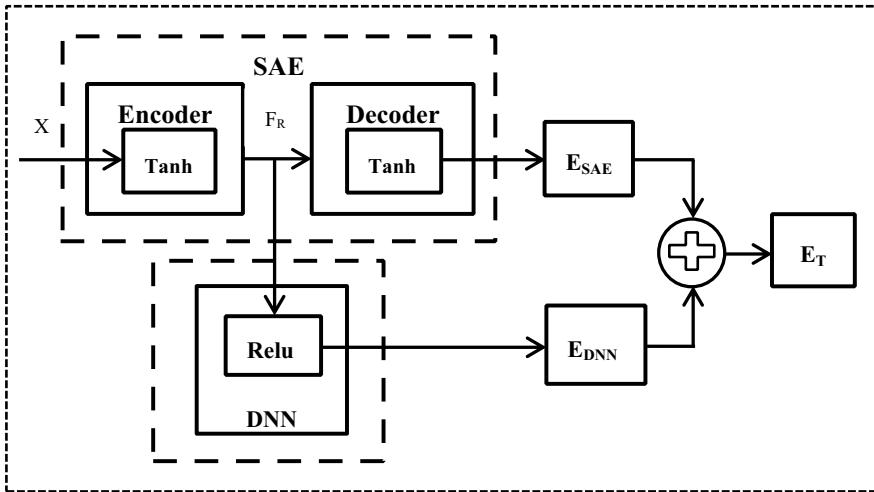
The framework consists of two different neural networks using two different activation functions for the prediction of diabetes (DB), hypertension (HT) and fatty liver (FL). The architecture of the proposed framework is shown in Fig. 1. The framework consists of three main components: An overcomplete SAE to learn and represent features from the original dataset in a new feature space, LP method for multilabel to multi class problem transform and deep neural network (DNN) architecture to extract high-level feature patterns from the improved feature space to predict the risk of multiple diseases. The first neural network used in the framework is the SAE and training of the SAE is done in using unsupervised learning method.  $L_1$  regularization is used as sparsity,  $L_2$  regularization is used as weight regularizer and Tanh is used as activation function in the SAE. Table 1 shows the configuration of the SAE. After training of SAE, only the encoder part of SAE is used in the final framework as shown in Fig. 2.



**Fig. 1** Architecture of the proposed framework

**Table 1** Configuration of SAE used in the framework

Parameters	Values
Number of nodes in the input layer	62
Number of nodes used as codes	100
$L_1$ regularization	0.001
$L_2$ regularization	0.001
Reconstruction error function	Mean square error



**Fig. 2** Total cost function of the proposed framework

Then in the next step LP transformation method is applied. In the last step, the second neural network that is the DNN architecture is used for the multiclass classification of the diseases. ReLU is used as the activation function in the layers of the DNN. The configuration of the DNN is shown in Table 2. Fine tuning of the framework is also done using supervised learning method on the training data by concatenating the two neural networks to minimize the classification error. Equation (1) shows the total cost function of the whole framework.

$$E_T = E_{SAE} + E_{DNN} \quad (1)$$

where Mean Square Error is used as error function in SAE and Categorical Cross Entropy is used as error function in DNN architecture.

**Table 2** Configuration of DNN architecture used in the framework

Parameters	Values
Number of nodes in the input layer	100
Number of hidden layers	3
Number of nodes in first hidden layer	256
Number of nodes in second hidden layer	128
Number of nodes in third hidden layer	32
Dropout	0.25
Batch size	128
Epochs	450
Loss function	Categorical cross entropy

## 4 Result Analysis

The dataset used for the study contains 110,300 entries collected for sixty-two features which include four basic physical examination items, twenty-six blood routine items, twenty liver function items and twelve urine routine items [8]. Presence of DB, HT and FL is predicted at the same time using the proposed framework. The dataset is divided into training and testing sets in the ratio of 80:20. The framework is implemented in python programming language. The results are compared on the basis of precision, recall, F1 score and accuracy. Accuracy achieved by using different activation functions in the framework is shown in Table 3. The performance of the proposed framework has been compared with other machine learning algorithms using LP transformation as shown in Table 4. Highest accuracy of 80.85% is achieved by using the Tanh activation function in SAE and ReLU activation function in DNN.

**Table 3** Accuracy achieved using different activation function in the framework

Activation function used in SAE	Activation function used in DNN	Accuracy (%)
Sigmoid	Sigmoid	51.77
Sigmoid	ReLU	53.05
ReLU	Sigmoid	57.84
ReLU	ReLU	58.75
Tanh	Sigmoid	75.71
Tanh	ReLU	80.85

**Table 4** Results achieved by different machine learning algorithms using LP transformation method

Algorithm	Precision (%)	Recall (%)	F <sub>1</sub> -score (%)	Accuracy (%)
SVM	87.16	72.34	78.37	67.78
KNN	73.93	61.38	66.54	52.67
Naïve bayes	78.05	62.51	68.93	55.56
Proposed framework	92.05	85.53	88.69	80.85

## 5 Conclusion

Prediction of multiple chronic diseases is a complicated process. More intelligent techniques are required to understand correlations among the attributes and label dependencies present in the multi-label dataset. A framework using SAE, LP problem transformation technique and deep learning architecture has been proposed in this paper. The framework performed better as compared to other machine learning techniques using LP transformation and achieved 80.85% of accuracy.

**Acknowledgements** The authors of this paper thank the Collaborative Innovation Center on Internet Healthcare and Health Service of Henan Province, Zhengzhou University, for providing the dataset, available online at <http://pinfish.cs.usm.edu/dnn/>.

## References

1. Garg A, Mago V, Role of machine learning in medical research: a survey. *Comput. Sci. Rev.* **40** (2021)
2. Cherman EA, Monard MC, Metz J, Multi-label problem transformation methods: a case study. *CLEI Electron. J.* **14** (2011)
3. J. Yu, X. Zheng, S. Wangb, A deep autoencoder feature learning method for process pattern recognition. *J. Process Control* **79**, 1–15 (2019)
4. D. Chartre, F. Chartre, S. García, M.J. del Jesus, F. Herrera, A practical tutorial on autoencoders for nonlinear feature fusion: taxonomy, models, software and guidelines. *Inf. Fusion* **78–96** (2018)
5. L.D. Mienye, Y. Sun, Z. Wang, Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. *Inf. Med. Unlocked* **18** (2020)
6. D. Preethi, N. Khare, Sparse auto encoder driven support vector regression based deep learning model for predicting network intrusions. *Peer-to-Peer Networking App.* **14**(4), 2419–2429 (2021)
7. V.J. Kadam, S.M. Jadhav, K. Vijayakumar, Breast cancer diagnosis using feature ensemble learning based on stacked sparse autoencoders and softmax regression. *Image Sign. Proc.* (2019)
8. R. Li, W. Liu, Y. Lin, H. Zhao, C. Zhang, An ensemble multilabel classification for disease risk prediction. *J. Healthc. Eng.* (2017)

# A Comparative Study of Different Machine Learning Based Feature Extraction Techniques in Bacterial Image Classification



Shallu Kotwal, Priya Rani, Tasleem Arif, and Jatinder Manhas

**Abstract** Bacteria are single-celled, microscopic organisms that have the tendency to survive in a variety of conditions. Majority of these organisms found in soil, ocean, and many of them are also present in human body. There are few species of bacteria that are beneficial to humans, whereas most of them are adverse in nature and are responsible for contamination that causes variety of infectious diseases that includes Cholera, Strep throat, Tuberculosis, etc. As a result, precise bacterial identification and classification is critical in preventing the spread of such life-threatening pandemic. Traditional methodologies-based bacteria image classification and identification techniques are time consuming, less accurate and needs extremely high skilled microbiologists to deal with complex nature of above said problem. With the evolution and penetration of machine learning based computer assisted technologies in this field, various flaws and issues can be addressed easily. The model developed using machine learning tools and technologies are highly successful in this domain of image analysis and have shown extremely high rate of improvement in clinical microbiology investigation by identifying different bacteria species. To improve upon results, feature extraction from digital images is essential and extremely important for better and accurate classification of bacteria. Feature extraction aids in removing unnecessary data from a data set and adds to the increase in speed of learning and generalization in entire machine learning process. This study presents a comparative study of research undertaken using different machine learning based techniques in feature extraction relating digital bacterial images that leads to effective and efficient classification and identification of different species. The study also identified and recommends the suitable classifier which is best at giving results using different feature extraction methods.

**Keywords** Bacteria · Microscopic · Machine learning · Digital images · Classification

---

S. Kotwal (✉) · T. Arif  
Department of Information Technology, BGSBU, Rajouri, India  
e-mail: [shallukotwal25@gmail.com](mailto:shallukotwal25@gmail.com)

P. Rani · J. Manhas  
Department of Computer Science and IT, University of Jammu, Jammu, India

## 1 Introduction

### 1.1 *Bacteria*

Bacteria are a type of prokaryotic and powerful microorganisms that have diverse role to play in earth's ecosystem. They come in a variety of sizes and shapes and usually measure between 0.2 and 20  $\mu\text{m}$ . Their structural examination needs an electron microscope [1]. Bacteria are beneficial to humans and environment e.g., those species which are employed in food fermentation, sewage treatment, fuel production, soil fertility maintenance, and pharmaceutical preparation. However, many bacteria species are hazardous to humans and other living organisms as they cause a variety of diseases like tetanus, food borne illness, typhoid fever, cholera, tuberculosis, etc. [2].

### 1.2 *Machine Learning*

Machine learning (ML) is considered to be a very important branch of artificial intelligence (AI). It can also be defined as “the area of computer science that enables and stimulate machines to learn without being explicitly programmed”. The domain like network packet classification [3], Prediction systems, sentimental analysis using NLP [4], real time speech recognition [5], disease diagnosis [6], financial industry, signal processing [7], fatigue analysis [8] are some of the applications that extensively employ ML techniques. Classification, regression, and clustering are three complex problems that ML can deal easily. The most extensively used ML algorithms are Linear Regression, SVM, Decision Tree, Naïve Bayes, ANN, RF, Deep Learning (DL), KNN etc. [9].

### 1.3 *Feature Extraction*

Feature extraction (FE) is a technique for reducing enormous amounts of redundant data. FE is the process of transforming raw data into a set of features. Extraction of features can be thought of as a pre-processing procedure that removes distracting inconsistencies from a dataset, allowing subsequent classifiers or regression estimators to perform better result. FE is an extremely important step in any kind of image classification framework [10].

## ***1.4 Types of Features***

The different kinds of features can be determined by system where the said problem is being implemented. Generally, features are broadly categorized as color, shape and texture.

### **1.4.1 Color Features**

Color is one of the most essential elements in image classification and its retrieval. Once the color space has been defined, the color features can then be recovered from the given set of images or regions. In this context, the most frequent approach is color histogram which involves extraction of colors during image processing [11]. Other methods like Color Coherence vector, Color Moments and color correlogram are also being used for the extraction of color features from different images.

### **1.4.2 Shape Features**

The features like shape are crucial for recognizing and distinguishing real world objects. The strategies for extracting shape features are divided into two categories: region and contour based. Contour based determines shape features from the boundary, while region-based pulls characteristics from the complete object [12].

### **1.4.3 Texture Features**

A texture is a collection of pixels with a unique appearance. Two types of texture feature extraction methods are available i.e. which involve (a) spatial texture features and another one for (b) spectral texture features. The first method extracts texture features by locating local pixel structures or computing pixel statistics in the original image, whereas the second method converts an image into the frequency and then calculates texture features from the modified image [12].

## **2 Different Features Extraction Techniques**

When important features are extracted, the model becomes more generic and less prone to overfitting. There are various FE techniques used by classification algorithms to predict outcome in machine learning.

## 2.1 Local Binary Pattern

Local binary pattern is a texture extraction technique that performs extraction of features in the form of textures [13]. The LBP involves converting the image into an array of integer labels thereafter calculating the intensity values for each cell ranging from 0 to 256. This is followed by calculation of intensity values of all of a cell's center pixels, number of comparable intensity values are counted, and the histogram is generated using count. Subsequently the obtained histogram is normalized, and the process is repeated for each and every cell in the image [14].

## 2.2 Gray Level Co-occurrence Matrix

GLCM is again texture-based feature extraction method. By applying second-order statistical procedures to the images, GLCM identifies the spatial textural relationship between pixels. GLCM Entropy, Correlation, Contrast, Energy and Homogeneity are five distinct texture properties [15].

“Entropy” =  $-\sum \sum q(i, j) \log q(i, j)$ ,  $q$  is the number of gray level co occurrence matrices in GLCM,

“Contrast” =  $\sum (i, j)2q(i, j)$ , here  $q(i, j)$  = pixel at location(i,j),

$$\text{Correlation} = \frac{\sum_{i=0}^{G_{\max}-1} \left\{ \sum_{j=0}^{G_{\max}-1} \{i, j, p_{i,j}\} \right\} - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y},$$

$$\text{Energy} = \sum_{i=1}^{G_{\max}} \{[h_i]^2\} \text{ and } \text{Homogeneity} = \sum_{i=1}^{G_{\max}} \left\{ \sum_{j=1}^{G_{\max}} \left\{ \frac{1}{1 + (i - j)^2} \cdot p_{i,j} \right\} \right\}$$

## 2.3 (HOG) Histogram of Oriented Gradient

The technique is again a feature extraction approach that calculates features from image after splitting it into blocks of  $2 * 2$  cells (we'll use  $4 * 4$  cells as an example), with each cell containing  $8 * 8$  pixels.

## 2.4 Deep Feature Extraction

Deep neural networks like CNN are distinct from traditional machine learning techniques. CNN retrieve features automatically rather than being provided manually. An input layer, multiple hidden layers, and a maiden output layer are typical components of deep neural networks.

## 3 Relevance of Feature Extraction in Image Classification

Extraction and selection of features is another milestone step which is calculated as the most relevant and extremely important in the process of classifying images. Most common features undertaken for consideration in bacterial image classification are shape, size, texture and color. In feature extraction process the researchers employ different extraction methods like HOG, Haralick features, Scale invariant feature transform, LBP, color histogram, CNN, etc. Feature extraction is relevant and plays a vital role in extracting the important information from the images under consideration. The said information will be further employed and subsequently used by the model in providing training to the system to perform efficient classification.

## 4 Literature Survey

An extensive literature survey has been done to collect the knowledge to use different machine learning methods employed in classification of different bacterial images. In Liu et al. [16] classified bacterial morphotypes considering morphological features. For classification, authors consider the KNN classifier and proposed Center for Microbial Ecology Image Analysis System called as “CMEIAS” an image analysis application for Windows NT. The proposed approach includes image segmentation applying threshold method, followed by extraction of size, shape and grey density based properties such as perimeter, roundness, area, major and minor axis length, and others. The authors employed the CMEIAS, an Image based Tool interface to activate measurement characteristics and then used varied measures of extracted features to create a tree classifier. The proposed method achieved 97% accuracy. Hiremath et al. [17] developed a cocci classifier tool to classify six types of bacterial species namely, *streptococci*, *cocci*, *diplococci*, *tetrad*, *staphylococcus*, and *sarcinae* by using ML techniques. The geometric shape features were extracted by labeling segmented region. With the use of extracted features, the three different ML algorithms namely K-NN,  $3\sigma$  and Artificial Neural Network are trained for classification. The accuracy achieved through KNN ranges from 91 to 100%, whereas with  $3\sigma$  classifier ranges from 92 to 99% and that with ANN classifier is 97 to 100%. In Rulaningtyas et al. [18] proposed ANN with back propagation algorithm for

the classification of tuberculosis bacterial images. Proposed approach involved suitable pattern recognition and classification technique in tuberculosis bacteria images by extracting geometric features, which include perimeter, area, radii, circularity, compactness, eccentricity and tortuosity. The classification is carried out using ANN with fine-tuned hyper-parameter values i.e. momentum with 0.9, mean square error with 0.00036, learning rate with 0.5 and number of hidden layers with 20. The given approach achieved with accuracy of 80%. Ayas et al. [19] developed Novel RF based technique for tuberculosis bacilli bacteria classification. The authors in this paper decreased the number of pixels manually and marked them erroneously in each image after comparing the RGB color components in the color space of each pixel with the color distributions. Noisy data elimination is also carried out using Mahalanobis distance. To propos RF-based classifier, the algorithm was trained on color pixels collected from different regions of the bacilli and non-bacilli bacteria. The method gave an accuracy of 89.38%. Turra et al. [20] implemented CNN based technique in classification of hyper spectral bacterial colonies. Proposed methodology involves spatial-spectral analysis and data normalization in extracting colony spectral signatures. The CNN architecture consisted of convolutional feature extraction layers alternated with pooling and a soft max layer in classification. The given methodology was compared with RF and SVM. This comparison shows a better overall performance. This approach achieved an accuracy of 99.7%, whereas the accuracy of SVM and RF was 99.5% and 93.8% respectively. Mohamed et al. [21] proposed a machine learning technique to classify ten bacterial species. Initially, the given images were converted from RGB to grayscale and finally enhanced by applying histogram equalization. After that, features are extracted by applying Bag of Words model, followed by use of SURF descriptors to extract local visual features in input images, and then classification was done by using multiclass Linear SVM. This proposed methodology achieved 97% of accuracy. Traore et al. [22] proposed and implemented CNN to classify image based data concerning Vibrio cholera and Plasmodium falciparum bacterial species. The proposed approach includes seven hidden layers which included 06 convolution layers, 01 fully connected layer and a soft max layer for final classification. Each convolutional feature extraction layer in this model is followed by ReLu activation function and max-pooling layer. The model shows 94% accuracy. Ahmed et al. [23] worked on Hybrid technique that were based on SVM and Inception-V3 techniques to classify microscopic bacterial images. The methodology involves image preprocessing that includes manual cropping of an image, conversion of image from grayscale to RGB, translation and flipping. Finally the feature extraction usingInception-V3 DCNN model is applied on the on preprocessed image. At the end the classification of the images is carried out using SVM. This method achieved accuracy with 96%. Mhatash et al. [24] proposed DL technique based methodology to classify 3D light sheet fluorescence microscopic images of larval zebra fish. The proposed method involved using the Gaussian method to separate and manually label the vibrio cholera images in three dimensional Zebra fish image species. Using this methodology the three-dimensional CNN was trained on the images undertaken for the study for features extraction and classification. The model employed sigmoid, Tanh and ReLu were activation functions to analyze its

accuracy. The said approach achieved accuracy of 95%. A comparison of several ML techniques has been performed and presented in Table 1. The most relevant papers that have been published by distinguished researchers in the domain of bacteria image classification have been included in Table 1. This table also provides a comparative analysis of the features, extraction techniques, classification algorithms, types of bacteria, dataset and the accuracy achieved by the given researchers during their experimentation.

## 5 Discussion

In this paper, the author presented a comparative analysis of ML techniques by using FE methods in bacterial classification. After going through the relevant literature, it has been discovered that researchers have successfully used feature extraction approaches in effective and efficient ML based bacteria image classification. Moreover, Feature extraction believes to be a crucial step in any kind of image classification. Processing any kind of image datasets always requires feature extraction technique. According to the literature available the shape, size, texture, and color are the most often used criteria for bacterial classification. After detailed review it is evident from the literature survey that different scientists working in this domain have heavily used different feature extraction methods such as simple geometric descriptor, Fourier descriptor, Surf descriptor and CNN feature maps to extract various features in any kind of bacterial image classification.

## 6 Conclusion and Future Scope

In this study, a comparison of different FE methods used for ML based bacteria image classification has been presented. These methods were applied on different types of dataset for feature extraction. The extracted features were given to different ML classifiers to classify and predict different bacteria species. The result in Table 1 concludes that the geometric and statistical feature and deep features have shown better result. Whereas the ANN classifier achieved higher accuracy as compared to other classifiers. Geometric and statistical features have shown a lot of potential in bacterial image classification. Different ML approaches, their objectives and the accuracy achieved by different proposed models have been thoroughly examined in this work. In future authors plan to study and apply more feature extraction algorithms on bacterial images for classification of bacteria. Based on different results, it can be inferred that a number of hybrid approaches can be developed to improve and enhance the accuracy of the system. In due course of time the authors intend to explore more feature extraction methods like feature fusion and feature attention mechanism to improve the classification result in digital microscopic images.

**Table 1** Comparative analysis based on different classifier

Classifier	Author/Year	Type of bacteria	Features	Extraction Method	Dataset used	Accuracy
RF	Ayas et al. [19]	Mycobacterium tuberculosis	Shape	RF	116 ZN stained sputum smear images	88.47%
	Turra et al. [20]	Hyperspectral bacterial images	Shape features	Shape features using PCA	16,642 Hyperspectral images	93.8%
KNN	Liu et al. [16]	Different bacteria species	Shape features	Shape features extracted from Fourier descriptor	4270 Digital images	97%
	Hiremath et al. [17]	Cocci bacteria cells	Geometric and statistical features	Geometric and statistical features using simple geometric descriptor	350 Digital microscopic images	91–100%
SVM	Turra et al. [20]	Hyperspectral bacterial images	Shape features	Shape features using Principal Component Analysis	16,642 Hyperspectral images	99.5%
	Mohamed et al. [21]	Bacteria species	Texture Features	SURF description	200 DIBaS dataset	97%
	Ahmed et al. [23]	Pathogenic bacteria	Texture and Shape features	Inception V3 model	800 HOWMED, PIXNIO	96%
ANN	Hiremath et al. [17]	Cocci bacteria cells	Geometric and statistical features	Geometric and statistical features using simple geometric descriptor	350 Digital Microscopic Images	97–100%
	Rulaningtyas et al. [18]	Tuberculosis bacteria	Shape features	Simple Geometric Features	100 ZN stained Sputum smear Images	Mean square error = 0.000368

(continued)

**Table 1** (continued)

Classifier	Author/Year	Type of bacteria	Features	Extraction Method	Dataset used	Accuracy
CNN	Turra et al. [20]	Hyperspectral bacterial images	CNN features	CNN method	16,642 Hyperspectral images	99.7%
	Traore et al. [22]	Vibrio cholera and Plasmodium falciparum	Deep features	CNN method	480 Digital microscopic images downloaded from Google	94%
	Mhathes h et al. [24]	Larval zebra fish intestine	Texture	CNN method	Zebrafish dataset	95%

## References

1. G.J. Tortora, B.R. Funke, C.L. Case, *Microbiology an introduction*, 10th edn. (CA Pearson Benjamin Cummings, San Francisco, 2010)
2. Bacteria, <https://en.wikipedia.org/wiki/Bacteria>, 23 Nov 2020(online)
3. B. Indira, K. Valarmathi, A perspective of the machine learning approach for the packet classification in the software defined network. *Intell. Autom. Soft Comput.* **26**(4), 795–805 (2020)
4. P. Kalaivaani, D.R. Thangarajan, Enhancing the classification accuracy in sentiment analysis with computational intelligence using joint sentiment topic detection with medlda. *Intell. Autom. Soft Comput.* **26**(1), 71–79 (2020)
5. Deng, L., Li, X., Machine learning paradigms for speech recognition: an overview. *IEEE Trans. Audio Speech Lang. Process.* **21**, 1060–1089 (2013)
6. I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective. *Artif. Intell. Med.* **23**(1), 89–109 (2001)
7. B. Nagaraj, D. Pelusi, J.I. Chen, Special section on emerging challenges in computational intelligence for signal processing applications. *Intell. Autom. Soft Comput.* **26**(4), 737–739 (2020)
8. K. Papagianni, M.A. Wahab, Multi-scale analysis of fretting fatigue in heterogeneous materials using computational homogenization. *Comput. Mater. Continua* **62**(1), 79–97 (2020)
9. R. Pillai, P. Qza, P. Sharma, *Review of Machine Learning Techniques in Health Care* (Springer, Proceeding of ICRC, 2019), pp. 103–111
10. A.O. Salau, S. Jain, Feature extraction: a survey of the types, techniques, applications, in *IEEE International Conference on Signal Processing and Communication (ICSC)* (2019), pp. 158–164
11. M.M. Sahu, A. Saxena, M. Manoria, Application of feature extraction technique: a review. *Int. J. Comput. Sci. Inf. Technol.* **6**(3), 3014–3016 (2015)
12. Dong ping Tian, A review on image feature extraction and representation techniques. *Int. J. Multimedia Ubiquitous Eng.* **8**(4), 385–395 (2013)
13. T. Ojala, M. Pietikainen, D. Harwood, A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* **29**(1), 51–59 (2019)
14. R.K. Gupta, N. Kumar, V. Sharma, M. Kaur, J. Manhas, Ensemble feature extraction-based detection of abnormal mass present in medical images using machine learning. *Adv. Intell. Syst. Comput.* **1187**, 241–251 (2020)
15. Muttag, W.K., S.K. Ali, Z.M. Aydam, B.H. Taher, Feature extraction methods: a review. *J. Phys. Conf. Ser.* **1591**, 012028 (2020)

16. J. Liu, F. Dazzo, O. Glagoleva, B. Yu, A.K. Jain, CMEIAS: a computer-aided system for the image analysis of bacterial morphotypes in microbial communities, in *Microbial Ecology Springer*, vol. 41 (2001), pp. 173–194
17. P. Hiremath, P. Bannigidad, Identification and classification of cocci bacterial cells in digital microscopic images. *Int. J. Comput. Biol. Drug Des.* **4**(3), 262–273 (2011)
18. R. Rulaningtyas, A.B. Suksmono, T.L.R. Mengko, Automatic classification of tuberculosis bacteria using neural network, in *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, Bandung (2011), pp. 1–4
19. S. Ayas, M. Ekinci, Random forest-based tuberculosis bacteria classification images of ZN-stained sputum smear samples. *Signal Image Video Proc.* **8**(s1), 49–61
20. G. Turra, S. Arrigoni, A. Signoroni, CNN-Based identification of hyperspectral bacterial signatures for digital microscopy, in *Lecture Notes in Image Analysis and Processing–ICAP*, Springer, vol. 10485 (2017)
21. B.A. Mohamed, H.M. Afify, Automated classification of Bacterial Images extracted from Digital Microscope via Bag of Words Model, in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*, Cairo, Egypt (2018), pp. 86–89
22. B. Traore, B. Kamsu-Foguem, F. Tangara, Deep convolution neural network for image recognition. *Eco. Inform.* **48**, 257–268 (2018)
23. T. Ahmed, M.F. Wahid, M.J. Hasan, Combining Deep Convolutional Neural Network with Support Vector Machine to Classify Microscopic Bacteria Images, in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh (2019), pp. 1–5
24. T.S.R. Mhatheesh, J. Andrew, K. Martin Sagayam, L. Henesey, A 3D convolutional neural network for bacterial image classification, in *Intelligence in Big Data Technologies Beyond the Hype*, vol. 1167 (2020), pp. 419–431

# Cloud Removal and Satellite Image Reconstruction Using Deep Learning Based Image Inpainting Approaches



Jaya Saxena, Anubha Jain, P. Radha Krishna, and Rajashree V. Bothale

**Abstract** The existence of clouds is one of the main factors that contribute incomplete information in optical remote sensing images, restricting their further monitoring applications of vegetation, land surfaces, water bodies, etc. Reconstruct the missing information caused by clouds can effectively improve the availability of remote sensing images. The problem can be broken into two parts detection & removal of clouds and reconstruction of the obscured data in satellite images. Several methods have been designed to solve this problem. Some require particular channels, while others are heavily dependent on the availability of temporally adjacent images. With the major breakthroughs in the field of deep learning and Convolution neural networks (CNN) attempts are being made for image reconstruction. In this work we use the image inpainting technique with two algorithms namely (Partial Convolutions and Generative Multi-column Convolutional Neural Networks) for the cloud removal in the satellite images. Image inpainting reconstructs an image with hole and fills the hole in such a way that it is not detectable. It maintains good global consistency with local fine textures. We evaluate our quality of result using SSIM and PSNR values and conclude that our approach is effective and efficient in removing small patches of thin clouds/haze.

**Keywords** Remote sensing · Cloud removal · Image inpaintaing

---

J. Saxena (✉) · R. V. Bothale  
NRSC, Hyderabad, Telengana, India  
e-mail: [jayasaxena@nrsc.gov.in](mailto:jayasaxena@nrsc.gov.in)

J. Saxena · A. Jain  
IIS (Deemed to be University), Jaipur, Rajasthan, India

P. R. Krishna  
NIT, Warangal, Telengana, India

## 1 Introduction

The presence of cloud is a significant barrier in optical-based remote sensing, limiting the use of satellite images for earth observation. Exploring the options for reconstructing/recovering data distorted by clouds or shadows is a critical and required pre-analysis phase. The existence and impact of cloud cover is governed by the factors like season, frequency of observation, and geographic location of the object. For earth observation purposes, clouded pixels are treated as no data values and should be filled. For cloud identification and reconstruction of contaminated pixels, a variety of techniques have been developed and deployed. Based on the geographical, temporal, and spectral structure of the data, these approaches can be divided into spatial, spectral and temporal categories. In the case of big irregular holes or missing sections, spatial-based methods such as interpolation and exemplar-based methods are unsuccessful. In the case of numerous corrupted bands, spectral-based approaches are ineffective. When significant changes occur in a short period of time or there is no data for the next period, temporal-based approaches such as Fourier transforms on time series data do not function effectively.

Thus, these traditional and classical approaches like Interpolation models, Wavelet transformation, Statistical methods, use of multi-temporal reference images etc. are applied for cloud removal and quality enhancement of reconstructed satellite images but they are restricted with many prerequisites and limitations. Much manual intervention and conventional hand crafted methods are required and are limited in terms of performance [1–3]. With the advent of newer technologies, deep learning and Convolutional Neural Networks (CNNs) proved major breakthroughs in the field of image classification and restoration [4–6]. The organization of the paper is as follows; Sect. 2 briefs about the related works in this area. We present our methodology and algorithms in Sect. 3. Section 4 presents quantitative analysis of the algorithms. Conclusion and future scope are presented in Sect. 5 followed by References.

## 2 Related Works

Presence of cloud is a very prominent and common problem in optical remote sensing. It adversely affects the quality of images resulting in information loss and restricts their applications for earth observation. The detection and removal of various types of clouds from satellite images have always been preprocessing steps for further analysis. A variety of methods and techniques were evolved and applied to address this problem. Many traditional approaches like time series reconstruction methods [3, 7, 8], usage of Filters [1], Fourier analysis [2], Wavelet transform etc. were applied however either they require many prerequisites or have limited applications with dependence on topology, resolutions and terrains. At times, it requires lot of manual intervention and expertise. With the advent of new technologies, Artificial neural networks and its successors, an era of data-centric approach evolved [5], giving a new

dimension to tackle this issue [6]. The major breakthrough in these fields gave rise to many algorithms and approaches [4, 9] for solving the task of image classification and reconstruction. The authors in [10] proposed modified Spatial Attentive Generative Adversarial Network for removing clouds in optical remote sensing.

In this work we use the image in-painting technique with two algorithms namely (Partial Convolutions and Generative Multi-column Convolutional Neural Networks) for the cloud removal in the satellite images. These algorithms are applied on a large date set to remove different types of clouds; thin clouds/haze and thick clouds/cloud shadows, and analyzed its performance. The output of our work was validated using visual observations and by PSNR, SSIM and MSE evaluation metrics. There is always a trade-off between performance and time taken or resources required. Thus, number of iterations for best optimum performance is also analyzed.

### 3 Proposed Image Inpainting Approaches

Many Deep learning algorithms which are suitable for image classification and restoration are studied. The two major algorithms in Image Inpainting, namely Partial Convolutions and Generative Multi-column Convolutional Neural Networks (GMCNN) are explored in detail. They are further applied and experimented using different data sets of various locations and varying cloud cover percentage.

It's the process of efficiently recreating missing sections to the point where it's impossible to tell which areas have been restored. This method is frequently used to eliminate undesired things from images or to restore damaged areas of old photographs. Image inpainting is a centuries-old technique that needed human painters to work by hand. However, academics have developed a number of automated inpainting methods in recent years. In addition to the image, most of these algorithms require a mask that shows the inpainting zones as input.

The modern era option is to inpaint images using deep learning and supervised image classification. The premise is that each image has a unique label, and neural networks are taught or “trained” to detect the mapping between images and labels over time. Deep networks show outstanding classification performance when trained on large training datasets (millions of photos with thousands of labels), and can often outperform humans. The basic workflow start by providing an input image with “holes” or “patches” that needs to be filled to the network. Because the network has no means of knowing what has to be filled in, these patches might be considered a hyper parameter required by the network. A photograph of a person without a face, for example, has no relevance to the network other than shifting pixel values.

We need a second layer mask that provides pixel information for the missing data in order for the neural network to comprehend what section of the image needs to be filled in. As it travels through the network layers, the input image undergoes numerous convolutions and deconvolutions. The network does create a completely synthetic image from scratch. We don't need to fill in the bits that are already visible

in the partial image, so we may use the layer mask to discard them. The output is then created by superimposing the newly generated image over the incomplete one.

### ***3.1 Partial Convolutions for Cloud Removal***

Initially the work started with the Image inpainting method based on partial convolutions for cloud removal tasks. To recreate the distorted image, existing deep learning-based picture inpainting methods use a standard convolutional network, with neural filter outputs conditioned on both valid pixels and the mean value in the masked gaps. Color disparity and blurriness are common artefacts as a result of this. Post-processing is commonly employed to minimise such artifacts, although it can be costly and ineffective. We propose using partial convolutions, in which the convolution is masked and renormalized so that it is only conditioned on genuine pixels. As part of the forward pass, we also incorporate a technique for automatically generating an updated mask for the following layer. We suggest using a Partial Convolutional Layer, which consists of a masked and re-normalized convolution operation followed by a mask-update step, to effectively handle irregular masks. When given a binary cloud mask, our convolutional results are solely dependent on the non-cloudy regions at each layer because we use partial convolutions. The automatic mask update phase is our key addition, and it removes any cloud masking that allowed partial convolution to function on an unmasked value. Even the largest veiled foggy parts will eventually shrink away with enough layers of successive updates, leaving only valid replies in the feature map.

#### **Implementation**

It uses a UNet-like architecture in the decoding stage, using nearest neighbour upsampling and replacing all convolution layers with partial convolution layers. The skip links will concatenate two feature maps and two masks, which will be utilised as feature and mask inputs for the succeeding partial convolution layer, respectively. The final partial convolution layer's input will be a concatenation of the original input image with cloud and the original cloud mask, which will allow the model to duplicate non-cloud pixels. The architecture is described in the following sections. In place of traditional padding, the partial convolutions are stacked with suitable masking at image boundaries. This ensures that erroneous values outside of the picture do not affect the inpainted content at the image border, which could be perceived as another hole. The keras framework is used to build the network, with Tensorflow serving as the backend framework.

#### **Dataset**

The dataset consists of  $512 \times 512$  size images of Sentinel 2 comprising around 1500 cloud-free images. The training folder consists of 1012 cloud-free images and the validation folder consists of 224 cloud-free images. The model is trained on binary cloud masks consisting of 366 cloud masks of  $512 \times 512$  size. Training.

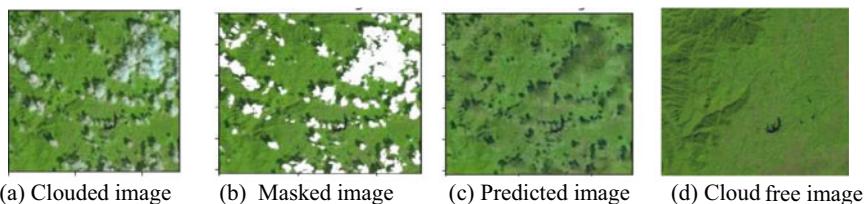
Network was trained on images of size  $512 \times 512$  from with a batch size of 8, and each epoch was specified to be 254 training batches long with total images of 1012 and 56 validation batches with images of 224. For training the model cloud free images along with binary cloud masks were provided. The training is divided into two phases. The first phase is trained with 50 epochs with batch normalization. In the second phase the network is trained with 50 epochs without batch normalization and a learning rate of 0.00005 because batch normalisation is a problem for masked convolutions. It takes around 4 h to train the model for each phase.

## Experiments

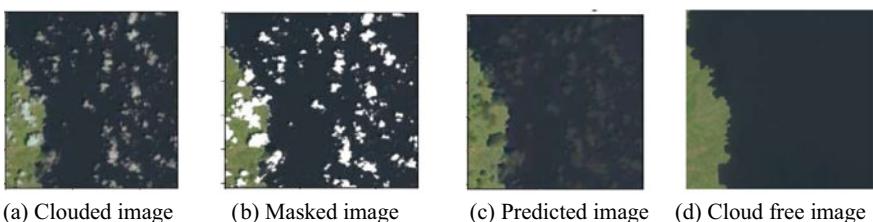
For evaluating the model's performance, in the first step the input clouded image is given to the model along with the binary cloud mask generated from the input image. For evaluating the model's performance, the model is tested on the Sentinel 2A, 2B satellite images consisting of RGB bands of JPEG format (Band 3,4,5), collected between the periods 2018 to 2020 with the varying cloud density of <40%. The input cloud images along with the binary cloud mask is provided to the model and the below images depicts the predictions made after training the model for 50 epochs without batch normalization. Data sets of three different locations are taken for experimentation, as shown in Figs. 1, 2 and 3.

## Observations

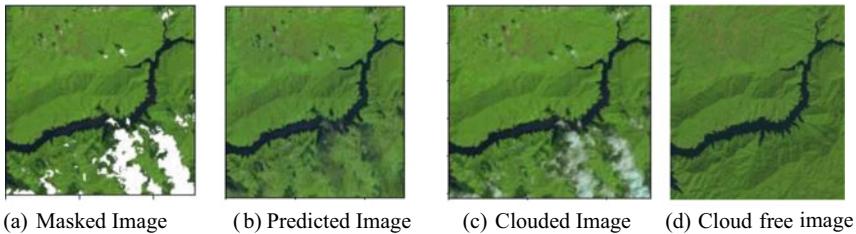
- This approach requires the cloud mask to be specified explicitly. The cloudy portion has to be identified in the clouded image and it has to be imposed with the cloud free image in order to reconstruct.



**Fig. 1** Austria latitude:  $47.5162^\circ$  N longitude:  $14.5501^\circ$  E, SSIM: 56% & PSNR: 25.7 dB



**Fig. 2** Sri Lanka latitude:  $7.8731^\circ$  N longitude:  $80.7718^\circ$  E, SSIM: 63% & PSNR: 27.3 dB



**Fig. 3** China latitude: 35.8617° N longitude: 104.1954° E, SSIM: 61% & PSNR: 26.7 dB

- Image inpainting technique using partial convolutions is not very effective for the reconstruction of the image after cloud removal. It fails to reconstruct the missing information present underneath the cloud. It just finds to substitute the missing information using the nearest neighbour pixels.

### 3.2 Generative Multicolumn Convolution Neural Networks

For image inpainting work, we have now explored a generative multi-column network. Within a single step, this network synthesises many visual components in simultaneously. We use a confidence-driven reconstruction loss to better characterise global structures. To improve local details, an implicit diversified MRF regularisation is utilised. To transfer local and global information gained from context to the target inpainting regions, a multi-column network with reconstruction and MRF loss is used. Picture inpainting (also known as image completion) is the process of estimating appropriate pixel information to fill in gaps in photographs. Using an image  $X$  and a binary region mask  $M$ , this inpainting approach may be learned from start to end (value 0 for known pixels and 1 otherwise). In image  $X$ , unknown parts are filled with zeros. Network is implemented using Keras Framework and using Tensorflow as backend of the framework.

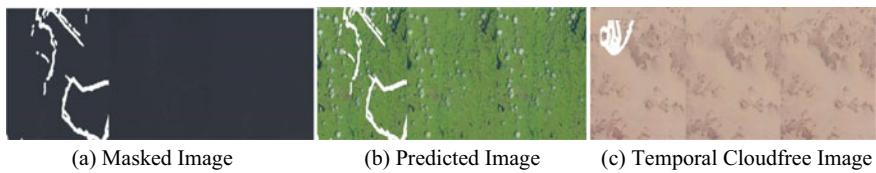
#### Dataset and Training

Model is trained on a dataset consisting of around 736 Sentinel RGB bands processed images of size 256X256 and on 110 masks from the mask dataset used for model training comes from NVIDIA's paper: Over 12,000 masks are included in Image Inpainting for Irregular Holes Using Partial Convolutions.

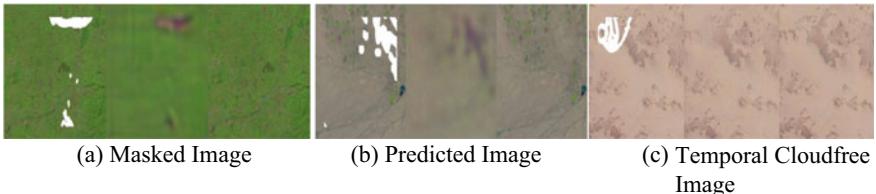
We have set the batch size as 1 and trained for around 110 epochs. Results of WarmUp Generator Training, is shown in Fig. 4 and Full Model training is shown in Fig. 5.

#### Experiments

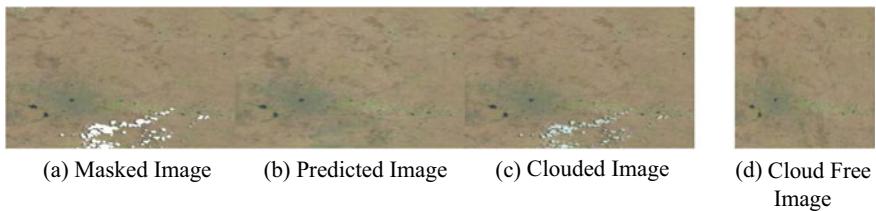
Data sets from three locations are taken for performing experiments, as shown in Figs. 6, 7 and 8. Sentinel 2A, 2B satellite images consisting of RGB bands (Band



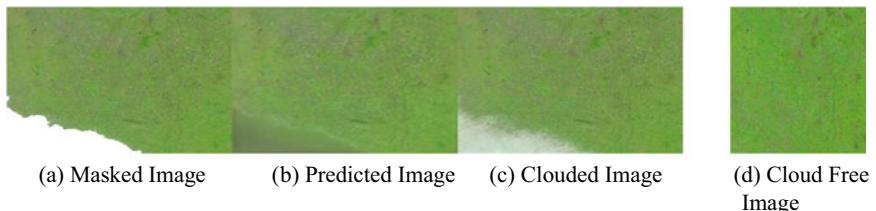
**Fig. 4** Results of warm-up generator training



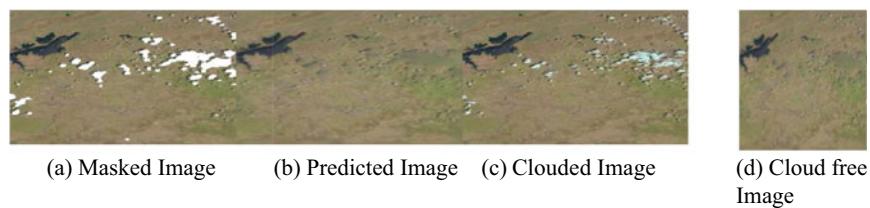
**Fig. 5** Results of Full Model training



**Fig. 6** Telangana, latitude:  $17^{\circ}23' 06''$  N longitude:  $78^{\circ}29' 12''$  E SSIM: 76% PSNR: 33 dB



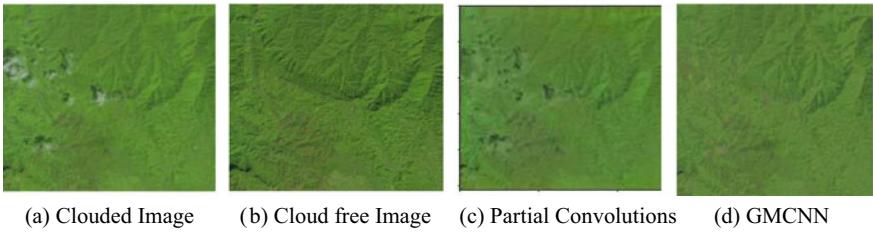
**Fig. 7** Paris, latitude:  $48^{\circ}51'23''$  N longitude:  $02^{\circ}21'07''$  E SSIM: 46% PSNR: 28.7 dB



**Fig. 8** Odisha, latitude:  $20^{\circ}57'06''$  N longitude:  $85^{\circ}05'54''$  E SSIM: 62% PSNR: 27.92 dB

**Table 1** Comparative analysis of approaches

Approach	Partial convolutions	GMCNN
PSNR, &SSIM	29.35 dB, 65.73%	28.47 dB, 60.65%



**Fig. 9** A sample comparative input and outcome

3,4,5) are collected between the periods 2018 to 2020 with the varying cloud density of < 40%.

### Observations

- This approach too requires the cloud mask to be specified explicitly. The cloudy portion has to be identified in the clouded image and it has to be imposed with the cloudy image in order to reconstruct.
- It is difficult to reconstruct the missing information lying under the cloud very efficiently and depends largely on the cloud cover percentage.

## 4 Quantitative Analysis

The evaluation metrics [11] of the Partial convolutions, GMCNN are compared and tabulated in Table 1.

A sample comparison of the output images obtained using these three approaches are presented in visual form, in Fig. 9.

## 5 Conclusions and Future Scope

Two algorithms of Image inpainting approach, namely Image Inpainting using Partial Convolutions and Image Inpainting using Generative Multi-column Convolutional neural Networks, have been proposed for the cloud removal (haze and thin clouds) from the satellite images. The inpainting approach requires the explicit generation of cloud masks from the input images. The cloudy portions have to be identified accurately in the input images. The generation of the cloud mask is a tedious task and it involves considering various variations of cloud like thickness, color, shape,

shadows. It is difficult to reconstruct the missing information lying under the cloud very efficiently and depends largely on the cloud cover percentage.

The model fails to achieve the promising results in case of the cloud cover density beyond 30%. Another disadvantage is that it has a tough time reconstructing images when there are dense or thick clouds. Also, when imaging conditions, such as lighting and the situation on the ground, change, satellite images acquired at different periods (even if they are of the same location) vary substantially in appearance, making it very difficult to obtain the ground truth of the area blocked by the cloud.

In future, it would be interesting to design an unsupervised method for this purpose. Furthermore, increasing the volume and variety of images can increase prediction accuracy in other areas. Other deep learning algorithms like GANs need to be explored for achieving better results. However, presence of thick clouds does necessitate the need of additional high wavelength imagery to gain some knowledge about the underlying ground information which involves future research and more multispectral bands into consideration.

## References

1. J. Chen, P. Jonsson, M. Tamura, Z.H. Gu, B. Matsushita, L. Eklundh, A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sens. Environ.* **91**, 332–344 (2004)
2. L. Olsson, L. Eklundh, Fourier-series for analysis of temporal sequences of satellite sensor imagery. *Int. J. Remote Sens.* **15**, 3735–3741 (1994)
3. T. Sakamoto, M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, H. Ohno, A crop phenology detection method using time-series MODIS data. *Remote Sens. Environ.* **96**, 366–374 (2005)
4. A. L'Heureux, K. Grolinger, H.F. Elyamany, M.A.M. Capretz: machine learning with big data: challenges and approaches. *IEEE Access* **5**, 7776–7797 (2017)
5. I. Good fellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press (2016)
6. J. Saxena, A. Jain, P. RadhaKrishna, A review on applicability of deep learning for remote sensing applications. *Solid State Technol.* **63**(6), 3823–3832 (2020)
7. J. Verbesselt, R. Hyndman, G. Newham, D. Culvenor, Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **114**(1), 106–115 (2010)
8. A. Verger, F. Baret, M. Weiss, S. Kandasamy, E. Vermote, The CACAO method for smoothing gap filling and characterizing seasonal anomalies in satellite time series. *IEEE Trans. Geosci. Remote Sens.* **51**(4), 1963–1972 (2013)
9. S. Mohajerani, T.A. Krammer, P. Saeedi, A cloud detection algorithm for remote sensing images using fully convolutional neural networks, in *IEEE 20th International Workshop on Multimedia Signal Processing (MMSP)* (2018), pp. 1–5
10. J. Saxena, A. Jain, P. RadhaKrishna, Modified spatial attention generative adversarial networks for cloud removal in optical remote sensing images. *Indian J. Natural Sci.* **12**(69), 35964–35974 (2021)
11. P. Ukrainski, Classification accuracy assessment. Confusion matrix method. *SCGIS Remote Sens.* (2016)

# A Novel Cipher Technique Using Substitution and Transposition Methods



Ritu, Niram, Ekta Narwal, and Sumeet Gill

**Abstract** The most fitting issue in the technological era is to provide security as most of the communications are done through the internet. Various classical encryption techniques such as Caesar cipher, Hill cipher, One-time pad, Rail-Fence techniques, and Columnar transpositions ciphers were designed to secure the data used in communication. These techniques are vulnerable to various active and passive attacks. To enhance security, we propose a new algorithm based on substitution and transposition techniques in which some bits are deleted and permuted in another place. The structure of ciphertext obtained through this algorithm is complex and thus, provides a high level of security. It can be used for the text of any length.

**Keywords** Cryptography · Encryption · Decryption · Substitution · Transposition

## 1 Introduction

In this technological era, network security is the most crucial issue in the world. Electronic frauds are increasing rapidly due to which organizations and individuals need to protect data and resources from unauthorized access. Thus, to secure networks and data, cryptography is being used. With the rapid growth of electronic communication, various encryption techniques such as cryptography, Steganography, digital signatures, watermarking, elliptic curve cryptography, and other applications are being used [1]. Cryptography is the craftsmanship and study of securing data. The specialty of securing data by changing into an indiscernible organization is known as cryptography. Cryptography comes from the Greek word ‘Kryptos,’ which means hidden and ‘graphy’ means writing [2]. In World War II, the Germans used the Enigma machine to encrypt the data for secret military communications. In the

---

Ritu · Niram · E. Narwal (✉) · S. Gill  
Department of Mathematics, M.D. University, Rohtak, Haryana, India  
e-mail: [ektanarwal.math@mdurohtak.ac.in](mailto:ektanarwal.math@mdurohtak.ac.in)

Ritu  
e-mail: [ritu.rs.maths@mdurohtak.ac.in](mailto:ritu.rs.maths@mdurohtak.ac.in)

nineteenth century, a popular hypothesis about the security rule of any encryption framework was proposed by Kerchoff [3]. This hypothesis has become the primary standard in planning a cryptosystem for analysts and specialists. The cryptographic algorithms provide confidentiality, integrity, availability, and authentication to users. They can be of three types depending on the keys and hash functions. Symmetric key cryptography, asymmetric key cryptography and hash function.

- (a) **Symmetric Key Cryptography:** The cryptographic technique in which only one key (known as a private key) is used to encrypt or decrypt the information, i.e., only one key is shared between sender and receiver. It is also known as Private Key Cryptography, e.g., DES, Triple DES, AES [4, 5].
- (b) **Asymmetric Key Cryptography:** The encryption techniques in which two keys (known as public and private keys) are used to encrypt or decrypt the information. It is also called Public Key Cryptography, e.g., RSA.
- (c) **Hash Function:** It is a trapdoor function that takes variable length size messages and produces a fixed-size output known as a hash value or message digest. It makes the content of plain text impossible to be recovered.

## 1.1 *Classical Encryption Techniques*

Symmetric key cryptography is the one of classical encryption technique which uses only one key for encoding and decoding of the data/information. The main feature of this type of encryption mechanism is, they use less computing power and also speed of data encryption/decryption is high in these [6]. There are two types of classical encryption techniques, which are as follows:

- (a) **Substitution Cipher:** It is a classical encryption technique in which each character in the plaintext is being replaced by other characters, numbers, or symbols. In this method, the characters are encrypted in a defined manner with the help of a key. The receiver decrypts the ciphertext by reversing the encryption process to obtain the original message. There are several kinds of substitution cipher such as Caesar cipher, Monoalphabetic cipher, Polyalphabetic cipher, Playfair cipher, Hill cipher, and One-time pad [7, 8].
- (b) **Transposition Cipher:** It is an encryption technique in which there is no replacement of letters, but the original message is being converted into ciphertext by performing some sort of permutation on the original message. It is just a rearrangement of the characters of the original message. A bijective function is used to encrypt the message and reverse the process to decrypt the ciphertext. Various transposition techniques are Rail- Fence Technique, Columnar Transposition cipher [9].

The rest of the paper is organized as follows: Sect. 2 presents a survey of related work, and the proposed algorithm will be presented in Sect. 3. Then, experimental results of the proposed algorithm will be given in Sect. 4. Finally, Sect. 5 provides the conclusion and future work.

## 2 Related Work

There are lots of works and studies present in cryptography using substitution and transposition methods. Researchers have used and are still using these two techniques in combination with other techniques for encryption and decryption processes. We have consulted several research articles to find out the relevant information about cryptography based on substitution and transposition methods. Vatsa et al. [10] proposed a novel cipher technique using the substitution method. In this method, the alphabets and numeric digits were being substituted with the help of a secret key. The proposed algorithm exploits the disadvantages of the existing symmetric substitution cipher technique. This technique was quite secure and efficient and resisted against Brute Force attack and Relative Frequency attack. It applied only to small pieces of text as it was difficult to keep track of position in a large text [10].

Rajput et al. [11] proposed an improved scheme in which the double encryption technique was used to encrypt the data securely. In the proposed scheme, Caesar's cipher technique with Hill cipher was used for encryption. The use of Hill cipher on encrypted data made it unstructured and resisted against Brute Force attack [11]. Poonia et al. [6] presented a paper in which a comparative study over classical encryption techniques was being studied. They included their study based on various parameters such as key size, type of key, and attacks on existing algorithms. A comparison table of various previous algorithms was presented. This was just to provide the basic knowledge of the existing algorithms [6]. Patel et al. [12] proposed a new substitution cipher called Random-X which could be used to encrypt the passwords. It was a unique substitution cipher that replaced each character of plaintext with triplets of characters. The main advantage of this technique was that the encrypted data corresponding to the same plaintext was not the same. It provided a high level of security and thus difficult to crack [12]. Purnama et al. [13] proposed a modified Caesar cipher method in which the cipher text could not be solved by a cryptanalyst. The proposed method replaced the characters into two parts, the vowels were being replaced with vowels too and the consonants were being replaced with consonants too. Some of the consonants in alphabets were not replaced. The main aim of this method was to conceal the message [13]. Orooba et al. [14] proposed a new algorithm based on substitution cipher and transposition cipher. The proposed method replaced the original message by deleting some bits and converting them into binary form and then placed them back in another place in plaintext. The two keys were used for encryption and decryption. This method could be used for both long and short text. The security could be increased by using steganography [14].

With the advancement of technology, the various existing algorithms are vulnerable to several attacks such as Brute Force attack and Relative Frequency attack. These algorithms thus do not provide security now a day. So, a new algorithm must be proposed based on classical encryption technique to provide security to the data and resist against these attacks.

### 3 The Proposed Algorithm

This algorithm includes a hybrid technique of substitution cipher and transposition cipher that makes it more secured and thus, difficult to crack. We are using two of the classical encryption techniques as a base to maintain the speed and efficiency of the encryption and decryption mechanism. This algorithm uses two keys, one is the length of plaintext, and the other is the length of deleted bits.

#### Algorithm for Encryption

**Step 1** Remove all spaces and convert it into a binary form using ASCII code.

**Step 2** Remove two bits after each 4th bit and store the string as key1, and remove bits as key2.

**Step 3** Divide key1 and key2 into two equal half parts and concatenate the string with 1st part of key2, after that 1st part of key1, 2nd part of key2 and 2nd part of key1.

**Step 4** Divide the string into two parts named I and II and concatenate the 2's complement of I and II.

**Step 5** Convert this binary form into hexadecimal form is the required ciphertext.

**Step 6** End.

#### Algorithm for Decryption

**Step 1** Calculate the length of key1 and key2 and store them as A and B.

**Step 2** Convert the given ciphertext into binary form.

**Step 3** Divide the binary form into two equal parts named I and II and concatenate the 2's complement of I and II.

**Step 4** First  $B/2$  bits are bits of key2, and the next  $A/2$  bits are bits of key1 and so on, and write the key1 and key2.

**Step 5** Insert two bits from key2 into key1 after every 4 bits and write the string.

**Step 6** Convert this binary form to text using ASCII code.

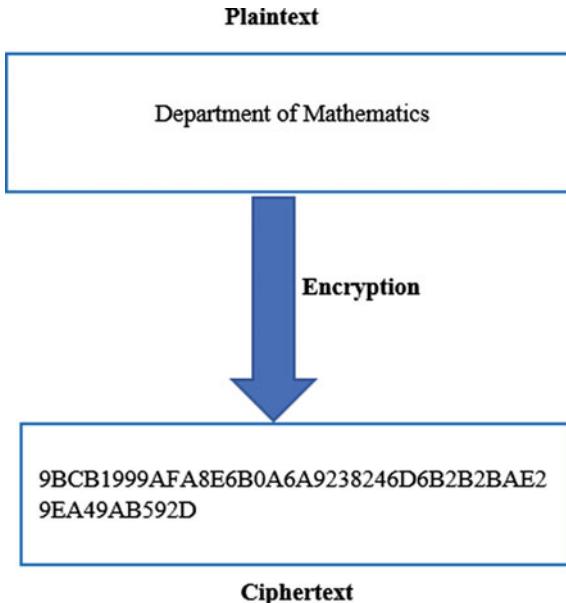
**Step 7** End.

### 4 Results

#### Results obtained from the Algorithms for Encryption and Decryption:

- **Encryption**

**Fig. 1** Plaintext to ciphertext



In Fig. 1, **Department of Mathematics** is the **plaintext** used for the encryption process. After applying various steps of the algorithm, we obtained ciphertext. **9BCB1999AFA8E6B0A6A9238246D6B2B2BAE29EA49AB592D** is the ciphertext, which is the result of first algorithm.

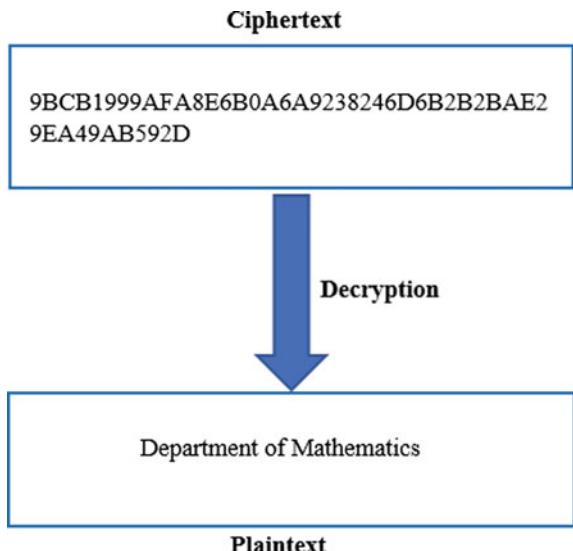
- **Decryption**

In Fig. 2, **9BCB1999AFA8E6B0A6A9238246D6B2B2BAE29EA49AB592D** is the ciphertext used for the decryption process. After applying various steps of the algorithm, we reobtained plaintext. The **Department of Mathematics** is the **plaintext**, which is the result of the second algorithm.

## 5 Conclusion and Future Work

This paper provides a novel algorithm based on substitution and transposition cipher by deleting some bits from strings and permuting them in another place. The algorithm is straightforward and can be implemented in all programming languages. The structure of ciphertext is complex and resists against Brute Force attack and Relative Frequency attack. It provides a high level of security and can be used for text of any length. In the future, we can also propose a new algorithm to make the structure more complex so that cryptanalysts cannot attack our message. Also, Steganography can be used instead of cryptography to increase security.

**Fig. 2** Ciphertext to plaintext



## References

1. E. Narwal, S. Gill, Simulating manual signature using Elman back propagation model to create pseudo digital signature. *Int. J. Innov. Technol. Explor. Eng.* **9**(2), 3548–3551 (2019). <https://doi.org/10.35940/ijitee.b6452.129219>
2. A.M., Qadir, N. Varol, A review paper on cryptography, in *7th International Symposium on Digital Forensics and Security (ISDFS 2019)* (October, 2019), pp. 1–6. <https://doi.org/10.1109/ISDFS.2019.8757514>
3. Menal, S. Gill, Performance evaluation of feed forward neural network for wired equivalent privacy/Wi-Fi protected access protocols. *Indian J. Sci. Technol.* **9**(31) (2016). <https://doi.org/10.17485/ijst/2016/v9i31/91847>
4. E. Agrawal, P.R. Pal, A secure and fast approach for encryption and decryption of message communication. *Int. J. Eng. Sci. Comput.* **7**(5), 11481–11485 (2017)
5. W. Stallings, *Network Security Essentials: Applications and Standards*, 4th edn. (Prentice Hall, New York, 2011)
6. P. Poonia, P. Kantha, Comparative study of various substitution and transposition encryption techniques. *Int. J. Comput. Appl.* **145**(10), 24–27 (2016). <https://doi.org/10.5120/ijca2016910783>
7. J. Abraham, S. Nanda, R. Gundla, Comparison between the types of substitute ciphers. *IJRECE* **7**(2), 67–71 (2019)
8. W. Diffie, M.E. Hellman, Privacy and authentication: an introduction to cryptography. *Proc. IEEE* **67**(3), 397–427 (1979)
9. A. Salai, T. Nadu, Evaluative study on substitution and transposition ciphers. *IJCRT* **6**(1), 155–160 (2018)
10. A.K. Vatsa, T. Mohan, S. Vatsa, Novel cipher technique using substitution method. *Int. J. Inf. Netw. Secur.* **1**(4) (2012). <https://doi.org/10.11591/ijins.v1i4.598>
11. Y. Rajput, D. Naik, C. Mane, An improved cryptographic technique to encrypt text using double encryption. *Int. J. Comput. Appl.* **86**(6), 24–28 (2014). <https://doi.org/10.5120/14990-2582>
12. F. Patel, M. Farik, A new substitution cipher—random-X. *Int. J. Sci. Technol. Res.* **4**(8), 125–128 (2015)

13. B. Purnama, A.H.H. Rohayani, A new modified caesar cipher cryptography method with legibleciphertext from a message to be encrypted. *Procedia Comput. Sci.* **59**, 195–204 (2015). <https://doi.org/10.1016/j.procs.2015.07.552>
14. O.I.I. Al-Farraji, New algorithm for encryption based on substitution cipher and transposition cipher. *Int. J. Curr. Res.* **7**(12), 23610–24612 (2015)

# Edge Computing: State of Art with Current Challenges and Future Opportunities



Vijay Singh Rathore, Vijeta Kumawat, B. Umamaheswari,  
and Priyanka Mitra

**Abstract** With the tremendous growth of IOT devices in ubiquitous fashion the data of such devices has increased enormously. By 2025, the world will produce 175 zettabytes (or 175 trillion gigabytes) of data. Traditionally such data are processed and stored in cloud storage. 91% of today's data is generated and stored in centralized data centre. Because cloud computing can't process large volume of data, the edge computing evolved. Edge computing also addresses some of the challenges of IOT like bandwidth, security, latency and noise ratio. More than 90 zettabytes of data will be produced by edge devices. By 2022, roughly 75% of all data would need analysis and intervention at the edge level. Such data can solve big challenges at a faster rate and allow business to access the benefit of edge opportunities for its growth. In this paper we will discuss about edge computing, its architecture, current challenges and future opportunities.

**Keywords** IOT devices · Cloud computing · Data centre · Edge computing · Architecture

## 1 Introduction

Edge computing is an open IT distributed architecture that enables Internet of Things (IoT) technology and mobile computing by processing data at the periphery of network by providing decentralized processing power. Rather than sending data to data centers, the processing is performed by local system or server in edge computing. Edge computing helps in improving the performance of Internet devices and web applications by performing computations closer to data source. The reduction in the

---

V. S. Rathore  
IIS (deemed to be) University, Jaipur, Rajasthan 302020, India

V. Kumawat · B. Umamaheswari (✉) · P. Mitra  
JECRC, Sitapura, Jaipur, Rajasthan 302022, India  
e-mail: [baluma78@gmail.com](mailto:baluma78@gmail.com)

long-distance communications reduces the need of latency and bandwidth between client and server.

### ***1.1 Comparisons Between Edge Computing and Cloud Computing***

Cloud Computing offers on-demand access to storage and computing devices via internet services in a centralized remote server network. Since the data is centrally maintained and processed virtually, it provides cost savings, mobility, and security to companies.

Edge Computing: For instance, it is the evolution of Cloud computing into a more powerful type. However, the literal sense states that all computation will take place at the network's edge rather than on a centrally run platform.

It takes processing closer to the end users by shifting it away from centralized servers. Edge computing is not introduced to replace cloud computing. It can be easily compared by linking them with SUV and racing car. Both SUV and racing car have different uses and purpose. Table 1 shows comparison to understand the difference between them.

## **2 Edge Computing Architecture**

### **2.1 Overview**

Until last decade the cloud has enabled many opportunities by processing and storing data in centralized data centers. But, as data has grown, the channel that hooks up that data to the cloud failed to work with. Also sometimes real time decision-making [1] has been delayed due to some problems like latency, bandwidth and compliance with the cloud. This paved an unprecedented way for edge computing to emerge and unlock new opportunities. The Edge architecture is nothing but a network architecture which process device data near the data source thereby started complementing cloud architecture (Figs. 1 and 2).

Device Layer: It consists of edge devices which is responsible for data acquisition, generation, transmission and delivery. Some edge devices provide limited gateway functionality and can perform some of the network functions. They are simple devices like sensors, smart watch etc. with little computing power and few resources. They generate data sets for heart beat, pulse rate, calorie burnt, temperature, humidity, stock prices, radiation [2] and more. The generated data are regulated by microcontrollers and must be sent to the Edge layer. Their communication and storage capacity [3] are also limited.

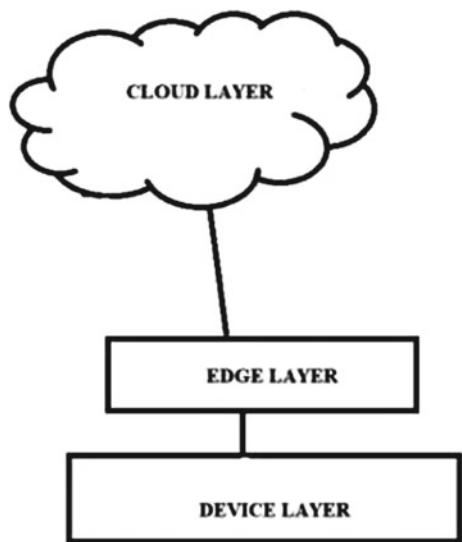
**Table 1** Edge computing versus cloud computing

Factors	Edge computing	Cloud computing
Suitable companies	For dealing with medium scale companies where budget limitations and latency concerns are important, edge computing can be used	For dealing with large data storage for organizations, cloud computing is more suitable
Programming	Different programming platforms can be used, all having different runtimes	Clouds require actual programming due to single target platform which uses single programming language
Security	Robust Security plan is required in edge computing which includes advanced authentication methods	Requirement of robust security plan is very less
Performance	Less Powerful in terms of processing	Powerful in terms of processing, cloud infrastructures are preferred over edge computing
Processing	Edge computing; only the initial processing is done by the “edge” network while rest is carried by a Centrally Managed System which makes the initial processing more time redundant and efficient	Cloud offers a centrally managed platform for the whole system of data retrieval or whatever the process might be, is initiated through the Centrally Managed System (CMS)
Data retrieval	Edge computing processing done on the “edge” network is through Internet of Things (IoT)	Data retrieval which includes applications and files can be accessed from the server directly in cloud computing
Applications	Edge computing is used for the real-time monitoring and analysis	Real-time analysis and monitoring cannot be efficiently performed through cloud computing

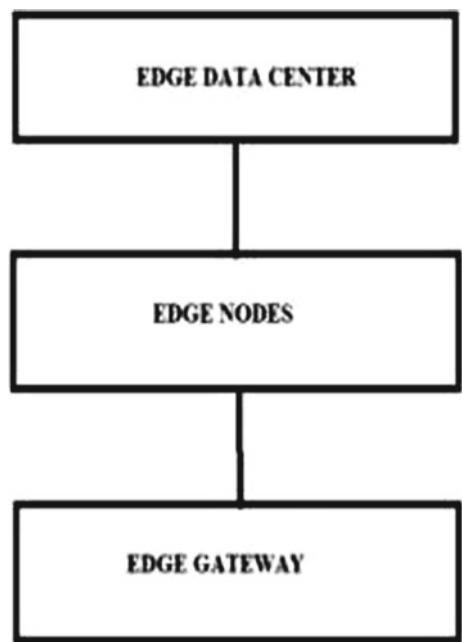
Edge Layer: It is the intermediate layer which is further divided into edge data-center, edge nodes and edge gateway. The edge data center carries the compute node and storage node. It stays near the edge. Edge node is a device or server or gateway that is responsible for performing edge computing. A gateway [3] is an edge cluster or server that is responsible for tunneling, firewall management, protocol translation, network termination and wireless connections and can host application workload.

Cloud Layer: This layer can be private or public cloud that is nothing but repository for different machine learning models. This layer is responsible for hosting and running different applications. The cloud manages different edge nodes by communicating with the workload [4] of edge nodes. Also cloud acts as a source and destination for various data that is accessed by other nodes.

**Fig. 1** Edge architecture



**Fig. 2** Detailed edge layer  
edge architecture consist of  
device layer, edge layer and  
cloud layer



### 3 Open Challenges

Edge Computing systems enables the task of execution within the device itself or outsource to local servers. Such systems minimize the data transmission but give rise to various technical challenges. Some of the important open challenges are as follows:

#### 3.1 Security

Hackers can easily access to the network core if there is any loophole in the edge security. The distinctive features of edge computing such as distributed architecture, location awareness, mobility support requirement, data processing are causing major challenge in ensuring security in edge computing paradigm. Moreover the edge devices designed are smaller in size than traditional server setups, which make them potentially vulnerable to physical manipulation. Thorough testing on the edge devices needs to be performed along with the understanding of security risks before adoption of this technology. There is a need to develop indispensible solutions for handling security issues with respect to access control systems, trust management, privacy, intrusion detection system, forensics and visualization. The edge devices should follow security protocols properly to ensure robust security. Artificial Intelligence and encryption techniques can be used to monitor and control potential threats. Machine Learning based security schemes are advisable for securing edge computing devices but many fatal defects are caused which are making such systems impractical to adopt.

#### 3.2 Cost

Configuration of an edge computing device and its deployment and maintenance are expensive. The older versions require additional equipment for processing which leads to extra cost. Moreover providing fault tolerance in edge computing system is highly costly as the backup server requires additional hardware along with high bandwidth. Machine Learning based predictive maintenance systems avoids unscheduled downtime and hence provides a cost effective solution. Also Machine Learning based security devices constitute high communication and computation cost due to large training data requirement for handling practical issues and processing functions. Thus there is a need to devise edge computing device with low communication and computation cost.

### ***3.3 Power***

High power processors, high voltage and three phase electricity are needed to provide customers and businesses cloud like remote service, which is a challenging task in remote areas.

### ***3.4 Data***

The subset of data is used by the edge computing devices for processing purposes such that sometimes additional data is there which is overlooked and sometimes raw data is available which gets wasted. Thus the challenge is to segregate data such that optimal data should be available to improve efficiency of the edge computing systems. Data sharing and computation across different edge devices are essential component to establish edge based distributed systems. Computation aware networking paradigm is a promising area for data sharing distributed systems.

### ***3.5 Heterogeneity***

The heterogeneity of data in a distributed system is an open challenge that needs to be resolved. Also the heterogeneous ecosystem of edge computing network makes the edge computing devices to access through multiple technologies such as 5G, 4G, 3G etc. Thus collaborations between different vendor systems is a challenging task. Factors such as load balancing, synchronization, heterogeneous resource sharing, data privacy, interoperability are making the collaboration challenging.

### ***3.6 Trust***

The trust of users on edge computing systems is closely related with the security and privacy of the system, which is itself a challenge. Thus the adoption and acceptance of this new technology by the user is based on the important factor of trust. Therefore developing consumer trust model that can stimulate the trust of a consumer for adoption of edge computing devices is an open challenge.

### **3.7 UseCases**

The functions should be clearly identified that can be performed in the cloud and those at the edges. This is a critical challenge for businesses as it clearly justify their investments.

### **3.8 Service**

Edge computing as a service helps the service providers and operators in fulfilling the demands of their customers with low infrastructure and revenue cost along with limited resources. The service discovery for increasing number of mobile users that require uninterrupted service is a challenging task. The delay in discovering and selection of available resources and services is another challenge for the edge computing systems.

## **4 Conclusion**

Till 2020, almost 45% of the world's data was stored and processed on the edge of the network, or perhaps even closer than this. Edge computing has witnessed huge growth in recent years compared to cloud computing. This paper surveys the edge computing technology in terms of overview, architecture, challenges and future scope. It also addresses the challenges related to IOT like security, cost, power etc. Providing solution to those problems is beyond the scope of this paper.

## **5 Future Scope**

This study provides researchers an important direction towards edge computing paradigm and its open challenges. Our future research will explore Edge based blockchain technology. The resource constrained edge devices produces data in a decentralized distributed setting which is highly vulnerable to security attacks. Edge based blockchain technology helps in prevention from security attacks and only sensitive data is transferred through decentralized edge infrastructure. But in resource constrained settings due to high energy requirements and computing power deployment of blockchain is very challenging [5].

- 5.1 Inés Sittón-Candanedo, Ricardo S. Alonso, Óscar García, Lilia Muñoz and Sara Rodríguez-González Edge Computing, IoT and Social Computing in Smart Energy Scenarios, <https://www.researchgate.net/publication/334>

- 816223\_Edge\_Computing\_IoT\_and\_Social\_Computing\_in\_Smart\_Energy\_Scenarios.
- 5.2 Edge computing architecture and usecases, <https://www.lfedge.org/2020/03/05/>.
- 5.3 **Xiaofei Wang, Yiwen Han, Victor CM Leung, Dusit Niyato, Xueqiang Yan, and XuChen, Convergence of edge computing and deep learning: A comprehensive survey. IEEE Communications Surveys & Tutorials, 2020.**

## References

1. R. Srinivasan, A. Zielinska, *Data at the Edge Managing and Activating Information in a Distributed World* (Seagate, 2019)
2. I. Sittón-Candanedo, R.S. Alonso, Ó. García, L. Muñoz, S. Rodríguez-González, Edge computing, IoT and social computing in smart energy scenarios [https://www.researchgate.net/publication/334816223\\_Edge\\_Computing\\_IoT\\_and\\_Social\\_Computing\\_in\\_Smart\\_Energy\\_Scenarios](https://www.researchgate.net/publication/334816223_Edge_Computing_IoT_and_Social_Computing_in_Smart_Energy_Scenarios)
3. M. Bischoff, J.M. Scheuermann, C. Kiesi, J. Hatzky, The edge is near: an introduction to edge computing, Gepostet June (2019)
4. Lfedge homepage, Edge computing architecture and usecases, <https://www.lfedge.org/2020/03/05/>
5. X. Wang, Y. Han, V.C.M. Leung, D. Niyato, X. Yan, X. Chen, Convergence of edge computing and deep learning: a comprehensive survey. IEEE Communications Surveys & Tutorials (2020)

# Deep Learning and Machine Learning Approaches for the Classification of Personality Traits: Review



Jayshri Patil and Jikitsha Sheth

**Abstract** Automatic personality traits assessment is an exciting area in affective computing. It received increasing attention in many computational areas, such as career counseling, relationship, and human resource management, forensic, recommendation system, and mental health diagnosis. The personality is an individual difference in thinking, feeling, and behaving pattern. The automatic recognition of personality traits involves understanding, prediction, and analysis of human behavior. The machine learning and deep learning approaches have been utilized by researchers to evaluate human personality. Deep learning is the most efficient, supervised, time and cost-efficient machine learning approach. This paper focuses on the comparative study of classical machine learning and deep learning approaches in personality classification tasks and their performance among all five traits.

**Keywords** Personality · Personality classification · Machine learning · Deep learning

## 1 Introduction

In psychology, personality is stable individual characteristics that are a combination of emotional, temperamental, and behavioral attributes. An individual personality influences their life choices, health, and preferences [1]. The prediction of preferences and behavior of people we found in all places in our day-to-day life. The personality recognition from the social network activity permits to predict preferences across situations and environments [2]. The recognition of user personality is useful in computational applications such as personality recognition would help psychologists to understand patient personality and its impact on patient behavior for

---

J. Patil (✉)  
BMCCA, Bhagwan Mahavir University, Surat, India  
e-mail: [jayshri.patil@bmusurat.ac.in](mailto:jayshri.patil@bmusurat.ac.in)

J. Sheth  
SRIMCA, Uka Tarsadia University, Bardoli, India  
e-mail: [jikitsha.sheth@utu.ac.in](mailto:jikitsha.sheth@utu.ac.in)

identifying personality disorders. In forensic, it is useful in analyzing conversations of suspected terrorists. User profiles with predicted personality traits would be used for the recommendation system. Personality trait affects one's suitability for certain jobs, in human resources management. By identifying personality traits in forensic help to reduce the circle of suspects.

## 2 Personality Measurements

The objective of psychologists is to understand human personality and its impact on human behavior and mental health. The personality is typically described in Big Five Personality Traits [3, 4]:

- **Extraversion**—Extraversion relates to sociability, talkativeness, and assertiveness. The people scoring high in Extraversion are more socially active, more friendly in public, express positive emotions, and show more agreements and complements than an introvert. Extraversion talks louder and more repeatability with fewer pauses and hesitation [4].
- **Openness**—The openness people appreciate new and uncommon ideas, and have a good sense of aesthetics. Openness is related to imagination, creativeness, curiosity, political liberalism, and appreciation for culture. Openness people prefer the longer words and tentative words as well as the avoidance of 1st person singular pronouns, and present tense forms in the conversation.
- **Agreeableness**—Agreeableness relates to maintaining positive social relations, being friendly, compassionate, and helpful. Agreeable people show compliance to satisfy others' demands even at cost of destroying their happiness.
- **Conscientiousness**—Conscientiousness people prefer an organized approach to life. Conscientiousness peoples are well organized, consistent, and reliable. Conscientious people do not prefer to use words related to negations, negative emotions, and discrepancies.
- **Neuroticism**—Neuroticism people experience mood swings and emotions such as guilt, anger, anxiety, and depression. Neurotics people are most likely to use 1st person singular pronouns, more negative emotion words, and less positive emotion words.

The Big-five factors of personality are important for accessing personality and also it is stable across the various studies.

## 3 Related Works

This section gives a comprehensive review of automatic personality recognition from the text based on the popular machine learning and deep learning models. In the paper [3], work is based on both verbal and nonverbal clues. The authors have been utilized

two gold-standard data sources such as the Stream-of-consciousness Essays dataset and the second source of data consists of conversation extracts recorded using an Electronically Activated Recorder (EAR). They have employed a total of four feature set LIWC, MRC, Utterance Type, and Prosodic features in their work. The authors have utilized classification and regression models for the personality classification task.

In this work [5] authors have been utilized commonsense knowledge with psycholinguistic features which gives a remarkable improvement in result. For extracting commonsense knowledge along with associated sentiment polarity and affective labels from text authors have employed the SenticNet dataset, ConceptNet, EmoSenticNet, and EmoSenticSpace. Authors have combined common sense knowledge-based features with psycho-linguistic features and frequency-based features (LIWC, MRC) and then the features were fed to the supervised classifiers.

The initial form of personality recognition systems [3–5] used a combination of machine learning algorithms and feature engineering. In recent years, a deep learningbased approach for personality recognition has been increased extremely. The deep learning approaches are used by the researcher in their experiments [6–9].

The approach in [9] utilized the deep learning architecture fully-connected (FC) network, convolutional networks (CNN), and recurrent networks (RNN) to automatically learn data representation. The experiment was performed on the myPersonality dataset and used the skip-gram method for word embedding. The results show that the best accuracy was achieved with the CNN with average pooling. The experiment in [7] was utilized an Essay dataset to extract user personality from it. The author used three Convolutional Filters to extract unigram, bigram, and trigram features from each sentence. The experiment was performed with the Convolutional Neural Network, author trained five different networks for the five personality traits as well as they utilized a two-layer perceptron. From the result, it is observed that the elimination of neutral sentences from the essay gives a marked improvement in the classification accuracy. The approach in [8] utilized Bidirectional Encoder Representation from Transformer (BERT) language model to extract contextualized word embeddings from textual data. The authors have employed a deep learning approach BERT word embedding with Bagged-SVM classifier for the personality recognition task.

## 4 Comparative Study of Automatic Personality Classification Based on Machine Learning and Deep Learning Algorithms

In the field of personality recognition, there is no recent work that provides the reader insights of an overall perspective of the advances in machine learning and deep learning-based automated personality recognition. This section presents a comparative study of machine learning and deep learning models for the personality classification task.

Machine learning is a branch of Artificial Intelligence that conveys to the system, the benefits of automatically learning from the concepts and knowledge. Deep learning is a sophisticated approach to machine learning and works on the Artificial Neural Network which learns from the data and makes an intelligent decision on its own whereas machine learning needs algorithms to understand data, learn from that data, and then produce accurate results. The detailed summary of machine learning and deep learning approaches utilized in the personality recognition task is shown in Table 1. It depicts the overall perceptive of the experiments performed using machine learning and deep learning approaches. Table 2 depicts the left to right, the number of samples involved in the experiments, performance on overall personality traits, and remarks.

Tables 1 and 2 infer that deep learning models bring the highest accuracy on a huge amount of data while machine learning requires a small amount of data to work and draw a conclusion. Deep learning architecture is built from many hidden layers and many neurons per layer. The multilayer architecture assists with the mapping of the input to higher-level representation. Deep learning creates new features on its own and whereas, in machine learning, features are accurately and precisely recognized by the users [10]. In experiments [3–6, 8–10], linguistic features have been utilized with the closed-vocabulary approach for the machine learning implementation. In the work [6–10], the implementation of deep learning utilized linguistic features of open vocabulary. The word embedding technique used for open vocabulary learns an internal representation of text. The results of the experiments [6–10] confirm that the implementation of open vocabulary with word embedding gives a better result of the personality classification task as compared to the machine learning algorithms with the closed-vocabulary approach (see Table 2).

## 5 Discussion and Conclusion

The paper represents the comparative study of machine learning and deep learning approaches used in personality recognition tasks. It has been concluded that the adoption of deep learning algorithms gave better accuracy than machine learning algorithms on the classification task. It easily extracts meaningful features, automatically learns the representation of data as well as captures temporal relationships between neighboring context words. Deep learning is a fast-growing approach to machine learning. The use of deep learning algorithms in the fields of personality recognition tasks shows its success and improves the accuracy rate in the results. Achievement with the use of Deep Learning models shows the growth of the deep learning approach for future innovation and research. There will be huge requirements of such approaches for implementing personality recognition tasks in industrial applications and for all human–computer interactions systems. The deep learning approach can also give successful results for other research domains such as digital image processing and speech recognition. Deep learning is the supervised

**Table 1** The summary of the experiments performed using machine learning and deep learning approach

**Table 2** The summary of the experiments using the machine learning and deep learning approach

Approach	Paper	Samples	EXT	NEU	AGR	CON	OPN	Remarks
Machine learning approach	[3]	2479 written essays	56.3 ACC	58.2 ACC	56.3 ACC	55.6 ACC	62.1 ACC	– Support vector machine and boosting algorithm produce higher classification accuracy – LIWC significantly performs better for all traits
	[3]	96 conversation transcripts	57.3 ACC	50.4 ACC	58.3 ACC	53.2 ACC	61.4 ACC	
	[6]	10,000 statuses of Facebook users	68.80 ACC	60.80 ACC	63.20 ACC	59.20 ACC	70.40 ACC	
	[6]	150 Facebook users	79.33 ACC	70.00 ACC	60.67 ACC	67.33 ACC	67.33 ACC	
	[5]	2400 written essays	0.634 F-score	0.637 F-score	0.615 F-score	0.633 F-score	0.661 F-score	– Inclusion of common sense knowledge features improved the performance of classifier
Deep learning approach	[6]	10,000 statuses of Facebook users	78.95 ACC	79.49 ACC	67.39 ACC	62.00 ACC	79.31 ACC	– Authors have concluded that the highest average accuracy is achieved with the implementation of deep learning models
	[6]	150 Facebook users	93.33 ACC	80.00 ACC	70.37 ACC	68.00 ACC	76.19 ACC	
	[7]	2467 written essays	58.09 ACC	59.38 ACC	56.71 ACC	57.30 ACC	62.68 ACC	– As the author in [6] proved that by utilizing a deep learning algorithm we achieve the highest accuracy, this is also confirmed by this experiment
	[9]	250 users with 9917 status updates	65.8 ACC	67.9 ACC	59.8 ACC	53.7 ACC	74.0 ACC	– The result of these experiments show that no improvement was achieved in n-gram CNN filters (continued)

**Table 2** (continued)

Approach	Paper	Samples	EXT	NEU	AGR	CON	OPN	Remarks
[8]	2479 written essays	59.30 ACC	50.39 ACC	56.52 ACC	57.84 ACC	62.09 ACC		<ul style="list-style-type: none"> <li>- As compared to context-independent Word2Vec utilized in [7], BERT gives the best representation of word and bagging improves classification accuracy</li> </ul>
[10]	2467 written essays	0.79 F-score	0.80 F-score	0.80 F-score	0.80 F-score	0.80 F-score		<ul style="list-style-type: none"> <li>- Utilization of deep learning pre-trained BERT model and classifiers improve performance</li> <li>- Combining word embedding with sentiment information improves the understanding of text content and accuracy</li> </ul>

and time-efficient machine learning approach. Because of its layered neural network architecture, it perceives tremendous growth in recent application domains.

Unlike traditional machine learning, deep learning techniques automatically generate features, but more requirement of a significant amount of labeled training data. For training deep learning models there will be a need for a huge amount of datasets for personality measures. We thus need to focus on how to collect sufficient amounts of data. In addition, we need to validate and clean the data for robust model training. Increased availability of resources and information about human behavior, the use of deep learning approaches in many research areas mainly human-machine interaction such as voice assistants, robots, cars, etc., and behavioral analytics will give significant progression and promising results. From the literature survey, it has been observed that a tremendous increase in the number of deep learning models in the last five years utilized for automated personality recognition tasks. The best performance in most personality recognition tasks has been achieved by deep neural networks. In the future, the implementation of deep learning algorithms will solve more complex functions and play a vital role in personality research and assessment.

## References

1. J. Patil, J. Sheth, Comparative study of data sources, features, and approaches for automatic personality classification from text. *Int. J. Comput. Appl.* **174** (2021). (0975-8887)
2. G.M. Vinciarelli, A survey of personality computing. *IEEE Trans. Affect. Comput.* **5**(3), 273–291 (2014)
3. F. Mairesse et al., Using linguistic cues for the automatic recognition of personality in conversation and text. *Artif. Intell. Res.* **30**, 457–500 (2007)
4. S. Han, H. Huang, Y. Tang, Knowledge of words: an interpretable approach for personality recognition from social media. *Knowl. Based Syst.* **194**, 105550 (2020)
5. S. Poria, A. Gelbukh, B. Agarwal, E. Cambria, N. Howard, Commonsense knowledge based personality recognition from text, in *Advances in Soft Computing and Its Applications* (Springer, Berlin), pp. 484–496
6. T. Tandera, Hendro, D. Suhartono, R. Wongso, Y.L. Prasetyo, Personality prediction system from Facebook users, in *2nd International Conference on Computer Science and Computational Intelligence*, Bali, Indonesia (2017), pp. 604–611.
7. N. Majumder, S. Poria, A. Gelbukh, E. Cambria, Deep learning base document modeling for personality detection from text. *IEEE Intell. Syst.* **32**(2), 7479 (2017)
8. J. Yu, K. Markov, Deep learning based personality recognition from Facebook status updates, in *IEEE 8th International Conference on Awareness Science and Technology (iCAST 2017)* (2017)
9. S. Dargan, M. Kumar, M.R. Ayyagari, G. Kumar, A survey of deep learning and its applications: a new paradigm to learning. *Arch. Comput. Methods Eng.* (2019). <https://doi.org/10.1007/s11831-019-09344-w>
10. J. Golbeck, C. Robles, M. Edmondson, K. Turner, Predicting personality from Twitter, in *Proceedings of IEEE International Conference on Social Computing* (2011), pp. 149–156

# Two-Level Priority Task Scheduling Algorithm for Real-Time IoT Based Storage Condition Assessment System



Saureng Kumar and S. C. Sharma

**Abstract** Most Internet of Things (IoT) tasks progressively run in real-time with an explicit deterministic timing requirement. Various real-time systems are available for IoT applications for task scheduling. A major drawback of this system is the real-time end-to-end transmission delay. To overcome the limitations of the existing system, We have proposed a hybrid algorithm with a combination of two algorithms first come first serve (FCFS) and a two-level priority-based task scheduling algorithm for a real-time IoT-based storage condition assessment system. This system has three sensor nodes Temperature and Humidity Sensor- DHT11, Air quality gas Sensor-MQ135, pH sensor Kit (E-201C-Blue) connected to a CC3200-LAUNCHXL microcontroller unit (MCU) with inbuilt wifi connectivity to perform real-time priority-based task scheduling according to their priority scheme. Our result shows that the proposed hybrid algorithm performs shortened transmission delay for priority tasks.

**Keywords** Internet of things · Task scheduling · Storage condition assessment · First come first serve

## 1 Introduction

A smart and intelligent system that sense efficiently communicates and transmits data and other devices is commonly known as the Internet of Things (IoT). Internet of Things can be applied in various applications like environmental monitoring applications, traffic monitoring, smart healthcare monitoring, and smart agriculture [1]. In this work, we proposed a real-time IoT-based storage condition assessment system that assesses fruit and vegetable freshness. There are three sensor nodes used to

---

S. Kumar (✉)

Cloud Computing and Wireless Sensor Lab, Indian Institute of Technology Roorkee, Saharanpur Campus, Saharanpur, India  
e-mail: [skumar@pp.iitr.ac.in](mailto:skumar@pp.iitr.ac.in)

S. C. Sharma

Electronics and Computer Discipline, Indian Institute of Technology Roorkee, Saharanpur Campus, Saharanpur, India

assess the freshness of fruits and vegetables, like humidity, temperature, pH, and air quality. This sensor is connected to the base station. We have set the temperature, and humidity is the highest priority subsequently. For priority-based scheduling, we have used a hybrid algorithm combination of first come first serve (FCFS) and priority-based task scheduling.

Task scheduling is a methodology of managing multiple tasks in a process model to minimize delay and optimize system performance. There are various algorithms like first come first serve (FCFS), round-robin (RR), virtual round-robin (VRR) are available for task scheduling. We undertook a comparative study of these algorithms based on their lowest waiting and turnaround time. We proposed a hybrid combination of two algorithms first come first serve (FCFS), and a two-level priority task.

According to the real-time scheduling analysis report by aviation research and development [2]. Real-time task scheduling can schedule the task within a specified deadline, depending on the characteristics of the scheduling algorithm [3]. To addressed stipulated time constraints for real-time packet scheduling, only a few researchers address end-to-end transmission delay. In this context, we tried to reduce the end-to-end transmission delay for the priority task. The rest of the paper is organized as follows. Section 2 present a literature study of the real-time priority-based scheduling task approach for IoT applications. Section 3 proposed a general overview of task scheduling algorithms and a comparative analysis of the different algorithms. Section 4 discusses the system platform for real-time IoT working, and Sect. 5 explains the proposed scheduling algorithm. Section 6 demonstrates the performance evaluation. Finally, the conclusion and future work are described in Sect. 7.

## 2 Literature Review

The priority-based task scheduling algorithm is the biggest challenge in real-time task scheduling because of several constraints of response time jitter, latency, time complexity, pre-emption, and energy consumption [4]. A suitable priority task scheduling must consider the priority of tasks because some task service requests processed earlier than the other task can't stay in the task queue for offloading. To address this problem, some intensive researchers widely investigated and proposed a new dynamic priority scheduling algorithm (DPSA) to solve the response time service request problem in the cloud computing system with a three-tier cloud structure (resource, service provider, consumer) [5]. Similarly, in IoT systems, It faces various attributes like task allocating problems that should be addressed for packet scheduling schemes. The most common task scheduling scheme uses first come first serve, which possesses overhead and transmission delay. To overcome this transmission delay author applied a three-class priority packet scheduling scheme in the field of wireless sensor networks [6] and save the network's scarce bandwidth and minimize energy consumption. In task scheduling, energy consumption is the biggest challenge. Several attempts have been made to overcome this issue. The author

proposed an enhanced genetic algorithm [7], ant colony optimization (ACO) algorithm [8], clonal selection algorithm [9] metaheuristic algorithm [10] that optimizing time, energy, makespan. Kavitha and Suseendran [11] introduced a priority-based adaptive scheduling algorithm (PASA) for IoT sensor systems to minimize data rate, delay power, energy, buffer size of the device, etc. The author compares his result with Energy Efficient Context-Aware Traffic Scheduling (EECATS) algorithm and finds the performance of the PASA algorithm in terms of packet delivery ratio (PDR), average residual energy of nodes, and throughput is better than that EECATS algorithm. But one of the significant challenges present in the IoT sensor system is to satisfy the quality of service (QoS) parameter while assessing resources to the task. The authors in [12] proposed a cost-aware scheduling algorithm that uses heuristic operations to assign tasks to the appropriate VM. The author claims that the experimental result demonstrates a high success rate and minimizes the execution cost under deadline constraints.

### 3 Task Scheduling Algorithm

#### 3.1 *First Come First Serve (FCFS)*

First come first serve (FCFS) algorithm is the most intuitive technique for real time computer systems, assigning processors to process the arrival task first. The next coming task receives services after the completion of the first arrived task. In this case, the real-time task might not be timely processed and cannot lead to a delayed time guarantee to network service [13].

The advantages of FCFS:

1. The scheduling algorithm is simple, easy to implement
2. FCFS provides predictable delay.

A maximum time delay occurs for the largest queue length, and a minimum delay occurs for the shortest. In the proposed algorithm, the highest count priority is executed first. The turnaround time (TAT) and waiting time are evaluated in similar priority FCFS. When two similar priorities are acting simultaneously, then the processor process which comes first is served first.

#### 3.2 *Round Robin (RR)*

Round robin algorithm is extensively used for real-time multitasking applications. It is the combination of pre-emptive and first come first serve scheduling. This restricted algorithm processor on the short time slice assigned to each processor. If the current process is not assigned within the stipulated time, then the processor executes the

next process, and the current process adds it to the last with the ready queue. As different queues have different group lengths, larger queue lengths get more service than smaller ones [14].

### 3.3 Virtual Round Robin (VRR)

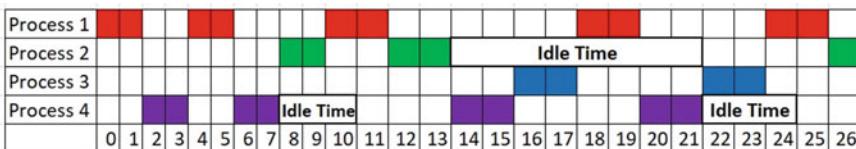
Virtual round-robin (VRR) algorithm is an optimal approach similar to the round-robin algorithm. In this algorithm, I/O bound process is released and shifted in the auxiliary queue hence. The auxiliary queue gets preference over the main ready queue. Since it has been selected for the main ready queue, it runs for quantum time minus total time spent running to avoid monopolization by the I/O bound process.

**Example:** Suppose we have 4 Processes, and their arrival time, CPU burst, and I/O burst are given as per the given in the Table 1.

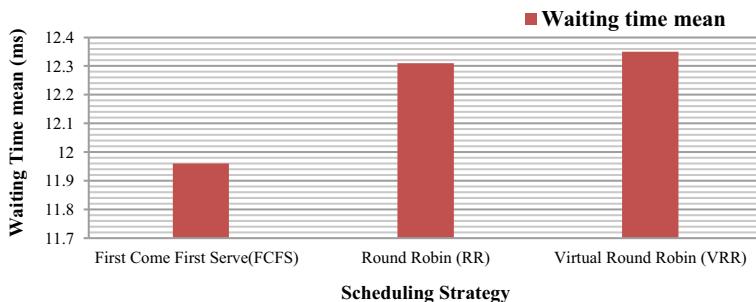
**Table 1** Process chart

Process name	Process 1	Process 2	Process 3	Process 4
Arrival time	0	5	12	2
CPU burst	14	4	5	4
I/O burst	7	8	13	3

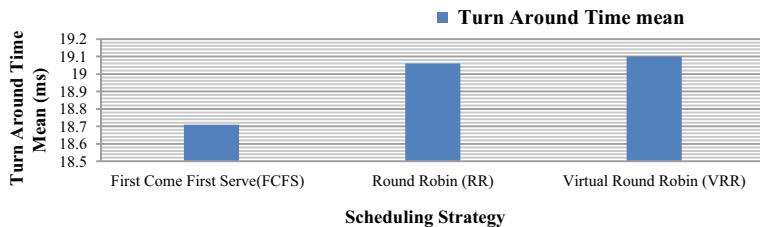
Note The above process chart are considered for comparison of all algorithm



**Fig. 1** Gantt chart for the process



**Fig. 2** Waiting time means for scheduling strategy



**Fig. 3** Turn around time mean for scheduling strategy

**Table 2** Comparison of scheduling strategy

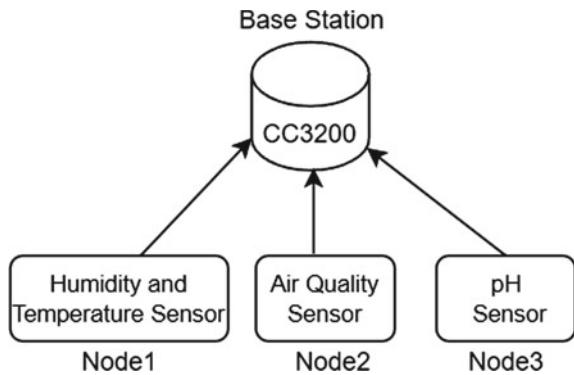
Scheduling strategy	First come first serve (FCFS)	Round robin (RR)	Virtual round robin (VRR)
Scheduling option	Time slice: 2	Time slice: 2	—
Process count	4	4	4
CUP busy time	200	197	197
CPU idle time	0	3	3
Processor utilization	100%	98.5%	98.5%
Fair schedule	No	Yes	Yes
TAT mean	18.71	19.06	19.1
TAT standard deviation	3.09	11.89	12.12
Waiting time mean	11.96	12.31	12.35
Waiting time standard deviation	5.24	7.72	7.94
Response time mean	11.99	3.41	3.41
Response time standard deviation	5.2	0.17	0.17

Although all the algorithms perform well, the speed of the process depends on end-to-end delay. In the first come first serve scheduling strategy, the mean waiting and turnaround times are the lowest (11.96 and 18.71 ms). Therefore, We proposed a hybrid algorithm with a two-level priority-based scheduling algorithm.

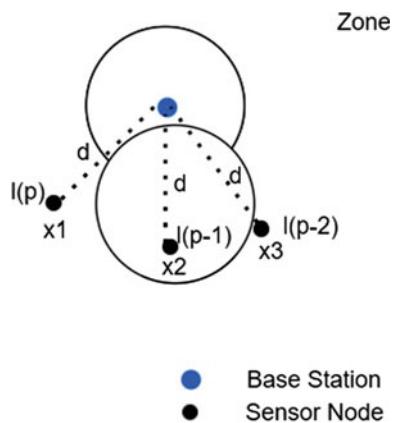
## 4 System Platform

Figure 4 Represent the architecture of IoT-based storage condition assessment devices connected with their processor to perform their task according to their priority. We have proposed three sensors Temperature and Humidity Sensor-

**Fig. 4** The architecture of IoT based storage condition assessment system



**Fig. 5** Node priority task scheduling



DHT11, MQ135—Air Quality Gas Sensor, pH Sensor Kit (E-201C-Blue) for assessment of storage conditions. This system is connected to a base station (CC3200-LAUNCHXL) with equal distance of sensor nodes. The CC3200-LAUNCHXL had a microcontroller (MCU) with built-in wifi connectivity. The task or packet scheduling is processed using the Time division multiple access (TDMA) scheme. Data packet sensed at the node and scheduled in the task queue. The base station allocates collision-free time slots for each sensor node based on their traffic priority. Now, we have set temperature as the first priority, and humidity is the second priority. The CC3200-LAUNCHXL is connected to the PC using the Energia software development program to assess the storage conditions (Fig. 5).

#### 4.1 Real-Time Task Scheduling System in IoT

The system performs three levels of scheduling approach.

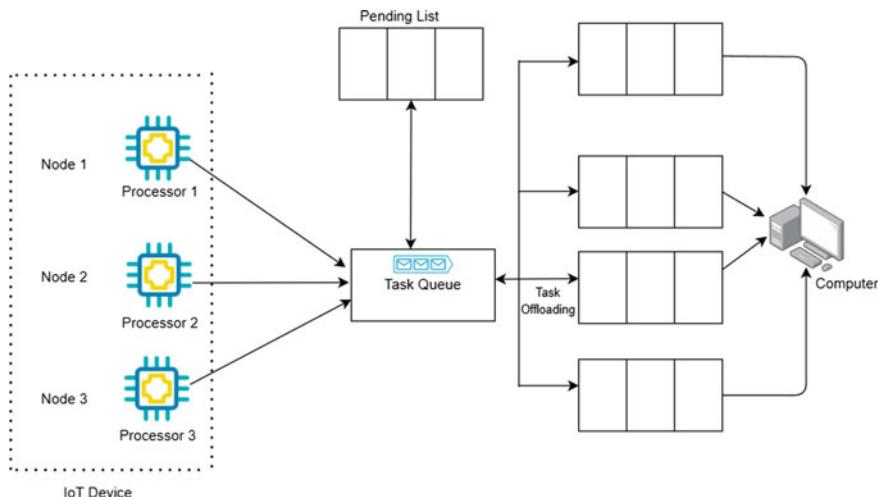
- Level 1:** is a set of tasks sent by IoT devices, which are required for execution.  
**Level 2:** is the task is created by IoT devices and executing using FCFS scheduling algorithm. It starts as pending whenever anyone criteria are met.

- If the task execution state changes to a non-terminal state (task queued, in progress)
- If old task execution changes into the terminal state (failed, timeout, reject, remove, cancel).

This level is responsible for tasks mapping with minimum completion time for all until the task is processed.

**Level 3:** is the task offloading. The entire sensor node is connected to the base station. It has inbuilt wifi connectivity for task offloading for further processing through the computer, as shown in Fig. 6.

The proposed system platform categorized the task according to their priority to process in IoT. There are two priority tasks, i.e., spatial priority and temporal priority tasks [15]. The spatial priority task needs information regarding the location, and the temporal priority task is time-specific. They give importance to the time of completion. Once the tasks arrive in the IoT gateway, first the processor will identify whether it is a spatial or temporal priority. In our research, we have considered spatial priority task scheduling algorithm to maximize resource utilization and minimize (execution time, waiting time, response time, memory, bandwidth, end to end transmission delay) for all task.



**Fig. 6** Task scheduling for IoT based storage condition assessment system

## 5 Proposed Scheduling Algorithm

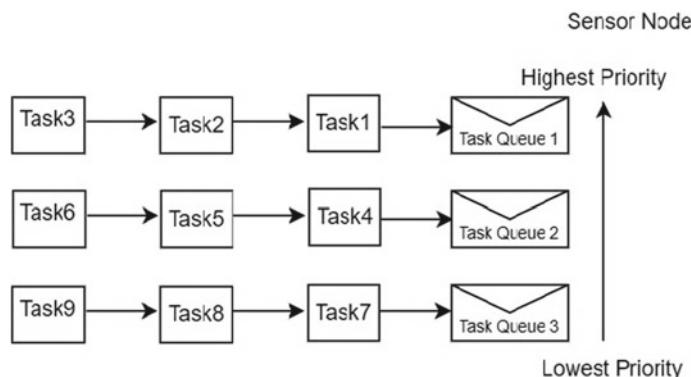
In this paper, we proposed a hybrid algorithm consisting of two first come first serve (FCFS) with two-level priority-based scheduling. Here, we have proposed a real-time IoT-based storage condition assessment system. This system comprises three sensor nodes followed by three-processor. The sensor node senses the data and the scheduling algorithm assign the processor to process the real time task, based on FCFS with two level priority, as shown in Fig. 7.

1. Set the N number of the node with its location
2. Each received data has a specific burst time required for the execution on the processor
3. Assign received priorities task  $P_1, P_2, \dots, P_n$  from the sensor node
4. Check and find out the highest top count priority task
5. Now calculate the transmission delay for the task.

**Algorithm** Two level priority task scheduling algorithm

```

while  $task_{p,i}$  sense data by  $node_i$  at level  $p$  i.e., at  $l(p)$  do
  if  $task_{p,i}$ , type = real-time then
    set  $task_{p,i}$ , in  $p1$  queue
  else if  $node_i$  is not at the lowest level
    if  $task_{p,i}$ , is not spatial then
      put  $task_{k,i}$ , in  $p2$  queue
    end if
    Assume time slot duration at  $l(p) \leftarrow t(p)$ 
    Task sensing time of  $node_i$  at  $l(p) \leftarrow senseTime_t(t)$ 
     $\therefore$  Residual time after task sensing,
     $t_1(p) \leftarrow t(p) senseTime_t(t)$ 
    Assume the total real-time task for  $node_i$  at  $l(p) \leftarrow n_k(p1)$ 
  
```



**Fig. 7** Task scheduling among multiple queues

```

Let,  $p1_{pro}(k) \leftarrow \sum_{j=1}^{n_k(p_1)} p1_{pro}(j)$ 
if  $p1_{pro}(k) < t_2(k)$  then
    All  $p1$  task of  $node_i$  at  $l(p)$  are executed by FCFS basis
    residual time  $t_2(k) \leftarrow t_1(k) - p1_{pro}(k)$ 
    Let,  $p2$  task of  $node_i$  at  $l(p) \leftarrow n_k(p_2)$ 
     $p2_{pro}(k) \leftarrow \sum_{j=1}^{n_k(p_2)} p2_{pro}(j)$ 
    if  $p2_{pro}(k) < t_2(k)$  then
        All the  $p2$  tasks are executed by FCFS basis
    else
        only  $p2$  tasks are executed for  $t_2(k)$  time
        no  $p3$  tasks are executed
    end if
    if  $p1$  task null &  $p2$  task executed at  $\alpha$ 
        consecutive timeslot since  $t_k \leq p2_{pro}(k)$  then
             $p2$  task are pre-empted at  $\alpha + 1, \alpha + 2, \alpha + 3 \dots \alpha + j$  timeslots
        if  $p1$  task take place during any of  $\alpha + 1, \alpha + 2, \alpha + 3 \dots \alpha + j$  timeslots
        then  $p2$  tasks are pre-empted  $p1$  tasks at any time and  $p1$  tasks are progressed
        Task transferred for  $p2$  task execution
    end if
    end if
end while

```

We have considered four-parameter (temperature, humidity, air quality, and pH) with two-level priority  $p1$  for temperature and  $p2$  humidity for real-time tasks. These sensor nodes required time to sense (*senseTime*) for the data packet at the instant  $t_k$  of real-time data  $node_i$ . Figure 5 shows the node  $x_1, x_2$  and  $x_3$  located at  $l(p), l(p - 1)$  and  $l(p - 2)$  and sensing the real-time data from these sensors. The node transmits priority 1 data, i.e., temperature data to the base station similarly, the node senses the humidity data and transfers to the base station. Suppose the priority processing time of real time at node1 is less than the real time data of node2. In that case, it can process the  $p1_{pro}$  data packet and if the maximum number of node  $n_m$  and maximum number of packets in the ready queue is  $p_m$  then the scheduling algorithm runs  $n_m \times p_m$ .

Now consider a scenario in which the real-time data reaches its neighboring node located at  $x_2$  the location. In the meantime, a second priority data, i.e., humidity data, is processed, transmission delay occurs from a node to the medium  $\frac{data_{p1}}{s_t}$ . The transmission delay from source to destination is calculated as  $\frac{d}{s_p}$ . Considering the same situation, the real-time data should satisfy the following inequality.

$$delay_{p1} \geq l(p) \times \left( \frac{data_{p1}}{s_t} + p1_{pro}(t) \right) + \frac{d}{s_p} + (l(p) \times t_{sq}) \quad (1)$$

The real-time data size represents data transmission speed. The distance between the base station and the node is  $d$ , where  $d$  is the medium length for wireless communication.  $p1_{pro}(t)$  is the processing time for priority 1 data at the node and  $t_{sq}$  is

the switching and queuing time. If priority 1 tasks ( $n_{p1}$ ) are in the queue, then the real-time task  $t_1$  has to wait. Now we consider the data size is the same for all tasks then,  $n_{p1}$  leading with  $t_1$  Hence, end to end delay becomes

$$\text{delay}_{t1} \geq \sum_{i=1}^{n_{p1}} (\text{delay}_{p1})_i \quad (2)$$

Similarly, for humidity (priority data 2)

$$\text{delay}_{p2} \geq l(p-1) \times \left( \frac{\text{data}_{p2}}{s_t} + p2_{pro}(t) \right) + \frac{d}{s_p} + (l(p-1) \times t_{sq}) \quad (3)$$

$$\text{delay}_{t2} \geq \sum_{i=1}^{n_{p2}} (\text{delay}_{p2})_i \quad (4)$$

Thus, by using the above equation and configuration variables, we have performed the priority task for end-to-end delay over the 4-node scheduling system.

## 6 Performance Evaluation

We have simulated the proposed algorithm using the C programming language and evaluate the performance against FCFS with two-level priority scheduling algorithms. In real-time, IoT performs fast data packet delivery. Nodes are uniformly separated from each other over the zone. Each node requires time to sense the data packet. This data packet is placed in the task queue based on their processing task. Moreover, we have set temperature as the highest priority, and humidity is the second-highest priority to process. We run the configuration variable over the number of the zone. Table 3 present the configuration variable for simulation.

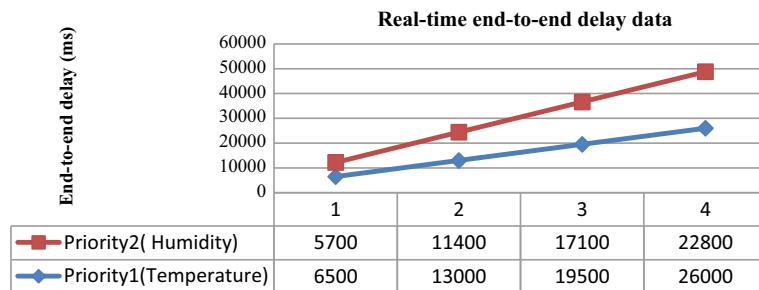
Figure 8 illustrate end-to-end delay of all real-time tasks over the number of the node. The proposed two-level priority temperature and humidity are used then the average transmission delay becomes shortened. The real-time priority task has lower transmission delay. This is due to the lower node task assigned with the higher priority of the current node. We would consider an expiration deadline data packet to minimize indirect computation time, memory, and bandwidth.

## 7 Conclusion and Future Work

In this paper, we have presented a comparative analysis study of first come first serve (FCFS), Round-robin (RR), Virtual round-robin (VRR) algorithms based on their mean waiting and turnaround time. We have proposed combination of two algorithms

**Table 3** Configuration variables

Configuration variables	Value
Switching time ( $t_{sq}$ )	10 $\mu\text{s}$
Propagation speed ( $s_p$ )	$198 \times 10^6 \text{ m/s}$
Data transmission speed ( $s_t$ )	250 Kbps
Distance (d)	100 (m)
Data size (for Priority 1)	8 B
Data size (for Priority 2)	7 B
Propagation speed for Priorities 1 and 2 over the wireless medium $p1_{pro}(t) = p2_{pro}(t)$	0.5 $\mu\text{s}$
Node 1 location ( $l_p$ )	0.2 m
Node 2 location ( $l_p - 1$ )	0.2 m

**Fig. 8** End to end real-time delay data over a number of zone

first come first serve (FCFS) with a two-level priority-based scheduling algorithm, in order to schedule the priority task queue for real time allocation with minimum makespan. This algorithm ensures real-time scheduling for IoT-based storage condition assessment systems. We have designed and evaluated the algorithm using C programming that performs efficient and optimal results in the transmission delay. Since we set temperature as the highest priority, humidity is the second-highest priority for real-time task scheduling. The data are processed with a minimum 6.5 and 5.7  $\mu\text{s}$  end-to-end delay subsequently.

In present study, security issues, network issues are not considered. We plan to consider data packet expiration time in task scheduling to further our research, ensuring minimum data packet expiration loss. Our finding might not consider the estimated processing time while selecting priority tasks. However, the time estimation process must be taken into consideration to prevent deadlock from occurring.

## References

1. M.S. Farooq, S. Riaz, A. Abid, T. Umer, Y.B. Zikria, Role of IoT technology in agriculture: a systematic literature review (2020), p. 41
2. U.S. Department of Transportation Federal Aviation Administration, Real-time scheduling analysis. Office of Aviation Research and Development, Washington, D.C. 20591, DOT/FAA/AR-05/27, Nov 2005 [Online] (2015). Available: [http://www.tc.faa.gov/its/wor ldpac/tech rpt/ar05-27.pdf](http://www.tc.faa.gov/its/worldpac/tech rpt/ar05-27.pdf)
3. P. Chennakesavula, J. Ebenezer, S.A.V.S. Murty, T. Jayakumar, Real-time packet scheduling for real-time wireless sensor networks, in *2013 3rd IEEE International Advance Computing Conference (IACC)*, Ghaziabad, Feb 2013, pp. 273–276. <http://doi.org/10.1109/IAdCC.2013.6514234>
4. L. Niu, W. Li, Energy-efficient fixed-priority scheduling for real-time systems based on threshold work-demand analysis, in *Proceedings of the Seventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis—CODES+ISSS '11*, Taipei, Taiwan (2011), p. 159. <http://doi.org/10.1145/2039370.2039397>
5. Z. Lee, Y. Wang, W. Zhou, A dynamic priority scheduling algorithm on service request scheduling in cloud computing, in *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, Harbin, Heilongjiang, China, Aug 2011, pp. 4665–4669. <http://doi.org/10.1109/EMEIT.2011.6024076>
6. L. Karim, N. Nasser, T. Taleb, A. Alqallaf, An efficient priority packet scheduling algorithm for wireless sensor network, in *2012 IEEE International Conference on Communications (ICC)*, Ottawa, ON, Canada, June 2012, pp. 334–338. <http://doi.org/10.1109/ICC.2012.6364150>
7. G. Shen, Y. Zhang, Power consumption constrained task scheduling using enhanced genetic algorithms, in *Evolutionary Based Solutions for Green Computing*, vol. 432, ed. by S.U. Khan, J. Kołodziej, J. Li, A.Y. Zomaya (Springer, Berlin, 2013), pp. 139–159. [http://doi.org/10.1007/978-3-642-30659-4\\_5](http://doi.org/10.1007/978-3-642-30659-4_5)
8. P.C. Pendharkar, An ant colony optimization heuristic for constrained task allocation problem. *J. Comput. Sci.* **7**, 37–47 (2015). <https://doi.org/10.1016/j.jocs.2015.01.001>
9. R.K. Jena, Energy efficient task scheduling in cloud environment. *Energy Procedia* **141**, 222–227 (2017). <https://doi.org/10.1016/j.egypro.2017.11.096>
10. M. Kumar, S.C. Sharma, S. Goel, S.K. Mishra, A. Husain, Autonomic cloud resource provisioning and scheduling using meta-heuristic algorithm. *Neural Comput. Appl.* **32**(24), 18285–18303 (2020). <https://doi.org/10.1007/s00521-020-04955-y>
11. K. Kavitha, G. Suseendran, Priority based adaptive scheduling algorithm for IoT sensor systems, in *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*, London, United Kingdom, Apr 2019, pp. 361–366. <http://doi.org/10.1109/ICACTM.2019.8776691>
12. X. Ma, H. Gao, H. Xu, M. Bian, An IoT-based task scheduling optimization scheme considering the deadline and cost-aware scientific workflow for cloud computing. *J. Wirel. Commun. Netw.* **2019**(1), 249 (2019). <https://doi.org/10.1186/s13638-019-1557-3>
13. W. Zhao, J.A. Stankovic, Performance analysis of 12 FCFS and improved FCFS scheduling algorithms for dynamic real-time computer systems, in *[1989] Proceedings. Real-Time Systems Symposium*, Santa Monica, CA, USA (1989), pp. 156–165. <http://doi.org/10.1109/REAL.1989.63566>
14. X. Li, P. Lun, Research and improvement of real-time queue scheduling algorithm, in *2010 International Forum on Information Technology and Applications*, Kunming, China, July 2010, pp. 102–104. <http://doi.org/10.1109/IFITA.2010.312>
15. U. Raza, A. Camerra, A.L. Murphy, T. Palpanas, G.P. Picco, Practical data prediction for real-world wireless sensor networks. *IEEE Trans. Knowl. Data Eng.* **27**(8), 2231–2244 (2015). <https://doi.org/10.1109/TKDE.2015.2411594>

# Intrusion Detection of Internet of Things Botnet Attacks Using Data Mining Technique



S. Kavitha, M. Hanumanthappa, and B. N. Kalavathi

**Abstract** In the current era, Intrusion Detection System maintains the system actions and store the log files in order to find the security issues. There is a necessity of improving the speed and accuracy in the network security in the network systems. We have applied clustering algorithms especially MakeDensityBasedClusterer, Canopy cluster, Kmeans cluster, FarthestFirst, Decision table. In this paper, we analyzed Botnet Attacks using Data Mining Techniques and the clustering algorithms for Intrusion Detection of Internet of Things. The farthestfirst cluster algorithm can be used for any anomaly detection to speed up the process as it took only 0.17 s to build the model.

**Keywords** Intrusion detection · Data mining · Clustering · IOT

## 1 Introduction

In the current era Intrusion Detection System maintains the system actions and store the log files in order to find the security issues. In the research area analysis of cluster is an major problem in the field of data analysis [2]. The important applications of clustering algorithms are like to group genes with similar expression patterns, identify malfunctioning servers. Mostly used clustering algorithms are K-Means [4] and Canopy cluster [3].

---

S. Kavitha (✉)

Assistant Professor, Dayananda Sagar College of Arts, Science and Commerce, Bangalore, India  
e-mail: [kavithas@dayanandasagar.edu](mailto:kavithas@dayanandasagar.edu)

M. Hanumanthappa

Professor, Department of MCA, Bangalore University, Bangalore, India

B. N. Kalavathi

Assistant Professor, Department of Computer Science, Dravidian University, Kuppam, India

## 2 Literature Survey

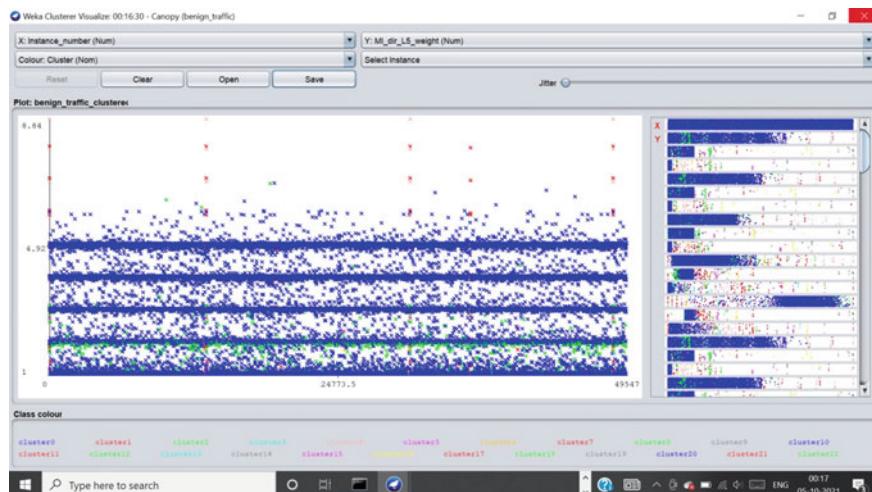
The paper Meidan et al. [1], authors proposed model which is applicable in extracting behavior snapshots of the network and also takes deep auto encoders to discovers anomalous network traffic from IoT devices [1] for anomaly detection. The paper Mirsky et al. [2] authors proposed a Kitsune: a plug and play NIDS to find detects in local network area in the online mode [2].

## 3 Data Set Description

The data set is collected online for the analysis of Internet of Things Botnet Attacks. Totally 7,062,606 instances and 115 attributes are there. Data Set characteristics are Multivariate and sequential [3, 4].

## 4 Results and Discussions

In the proposed method, we aimed to distinctive between two types of anomaly detection technique benign and Malicious traffic data by anomaly detection methods. In this paper used clustering algorithms such as Canopy cluster, K-means cluster,



**Fig. 1** Canopy cluster

MakeDensityBasedClusterer, Farthestfirst and Decisionstump algorithms and the results are shown below (Fig. 1).

### **Canopy Cluster**

Based on the based on attribute standard deviation value, canopy cluster can be used in batch mode, for set the T2 distance value. The value of T2 value calculates how many clusters are created. Canopy cluster produces the output of first N clusters if the user mentions the N number of canopies to generate where N is less than number of clusters. In the canopy cluster will use the single pass execution on the data. Canopy cluster can execute in two ways either in batch mode or the other one is incremental mode.

Test mode: evaluate on training data

Canopy clustering

Number of canopies (cluster centers) found: 23

T2 radius: 1.689

T1 radius: 2.111

Time taken to build model (full training data): 3.26 s.

### **K-Means Clustering**

The collected data is belongs to undefined category, then we can apply  $K$ -means clustering. The importance of the K-Means clustering algorithm is to discover set of similar group in the data and the number of similar groups represented  $K$  variable. When features are added each time as an input, the job of this algorithm is to assign each data point to one of  $K$  groups. Figure 2 shows the resultant of KMeans clustering on the given dataset.

The result of K-Means cluster is as shown below

Number of iterations: 5

Within cluster sum of squared errors: 12,378.963542740934

Missing values globally replaced with mean/mode

Time taken to build model (full training data): 0.78 s.

Clustered Instances

0 18,474 (37%)

1 31,074 (63%)

### **MakeDensityBasedClusterer**

Figure 3 shows the resultant of MakeDensityBased clustering on the given dataset.

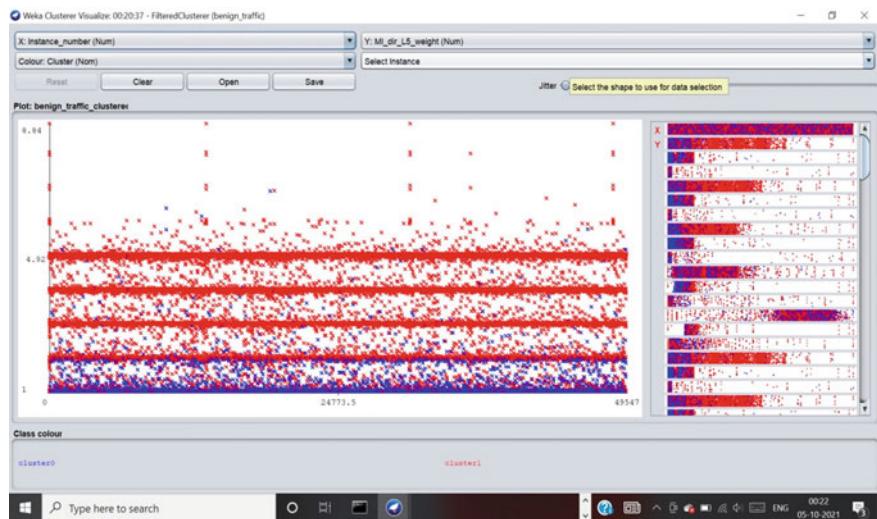
Within cluster sum of squared errors: 12,378.963542740934

Cluster: 0 Prior probability: 0.3729

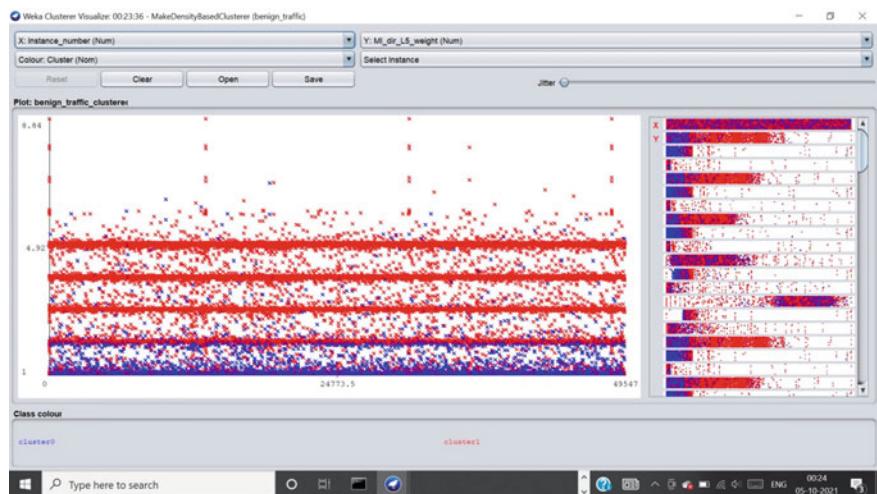
Time taken to build model (full training data): 0.83 s

Clustered Instances

0 18,776 (38%)



**Fig. 2** K-means clustering



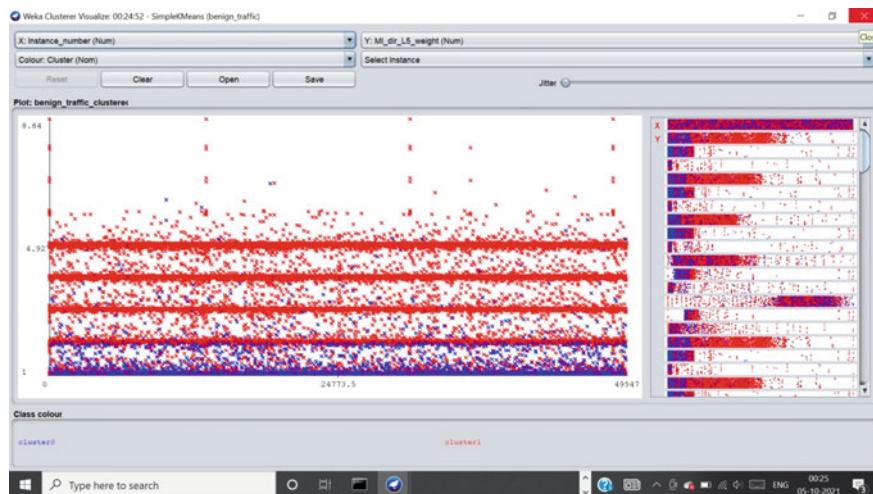
**Fig. 3** MakeDensityBasedClusterer

1 30,772 (62%)

Log likelihood: -595.20113.

### kMeans

Figure 4 shows the resultant of kMeans clustering on the given dataset.



**Fig. 4** kMeans

Number of iterations: 5

Within cluster sum of squared errors: 12,378.963542740934

Time taken to build model (full training data): 0.64 s.

Clustered Instances

0 18,474 (37%)

1 31,074 (63%)

### FarthestFirst

In the farthest-first traversal the first point is chosen arbitrarily and each following point is as far as possible from the set of previously-selected points. All the preceding point of a farthest-first traversal gives a set of points that is extensively spaced and near to all remaining points. Figure 5 shows the resultant of FarthestFirst clustering on the given dataset.

Time taken to build model (full training data): 0.17 s.

Clustered Instances

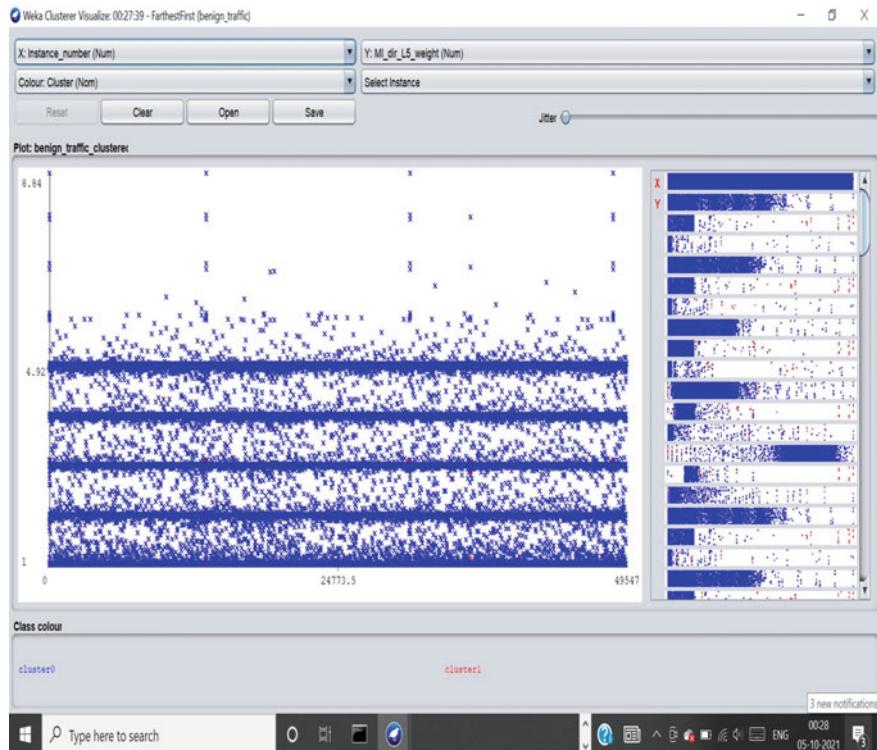
0 49,515 (100%)

1 33 (0%)

### Decision Table

When multiple conditions to be applied on the dataset then we can use the Decision table clustering method. The decision table is very much useful in recording the various actions and decisions to be taken in group of conditions.

Number of training instances: 49,548.



**Fig. 5** FarthestFirst

Number of Rules: 166.

Non matches covered by Majority class.

Best first

Start set: no attributes

Search direction: forward

Stale search after 5 node expansions

Total number of subsets evaluated: 1098

Merit of best subset found: 0.004

Evaluation (for feature selection): CV (leave one out)

Feature set: 59, 61, 65, 80, 111, 115

Time taken to build model: 43.78 s.

## 5 Conclusion

Anomaly detection is having prominent role in big data analytics. It is having applications like behavioral analysis in order to aid in learning about the detection, identification and prediction of the occurrence of these anomalies. Intrusion detection plays an important role in network security system. The paper is concentrated on the comparison of clustering methods for intrusion detection dataset using data mining tool. The study is concluded that the K-means clustering given the accurate results.

Even though many research is going in anomaly detection, in our research paper, have given the classification algorithm which is used less time complexity. The K-Means cluster algorithm can be used for any anomaly detection to speed up the process as it took only 0.02 s to build the model.

## References

1. Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai, Y. Elovici, N-BalIoT: network-based detection of IoT botnet attacks using deep autoencoders, in *IEEE Pervasive Computing, Special Issue—Securing the IoT* (2018)
2. Y. Mirsky, T. Doitshman, Y. Elovici, A. Shabtai, Kitsune: an ensemble of autoencoders for online network intrusion detection, in *Network and Distributed System Security (NDSS) Symposium*, San Diego, CA, USA (2018)
3. Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai, Y. Elovici, dataset
4. Meidan, Bohadana, Mathov, Mirsky, Shabtai, Department of Software and Information Systems Engineering; Ben-Gurion University of the Negev; Beer-Sheva, 8410501, Israel

# Code Tamper-Proofing Using Return Oriented Programming in IoT Devices



Rajesh Shrivastava , Simar Preet Singh ,  
and Mohammad Kamrul Hasan

**Abstract** An application running on the Internet of Things (IoT) device is vulnerable to code reuse attacks, which makes control flow highjacking relatively easy. Control Flow Integrity (CFI) is widely known to make code tampering harder. Available CFI solutions mostly work for x86 (Intel architecture), but most of the IoT devices use the ARM architecture. This paper presents a novel control-flow obfuscation technique for IoT devices, especially for the ARM architecture. This paper uses Return Oriented Programming (ROP), used for code reuse attacks, to prevent control flow in IoT devices. The proposed method is a self-contained code integrity verification approach that protects control flow by ROP gadgets and implicitly verifies integrity by “U-probe”.

**Keywords** Control flow integrity · ARM architecture · Return oriented programming · Code tampering · Internet of Things (IoT) security

## 1 Introduction

The Internet of Things (IoT) is a distributed network that consists of a large number of physical devices called IoT nodes. IoT deployment involved in various applications, such as health care, agriculture and smart home. According to a survey on Industrial IoT Security Survey by SANS [1], 72% of IoT nodes connected directly to Internet and bypassing traditional security layer. Additionally, as many IoT devices are left unattended and automatically collect data (possibly in hostile environments) and

---

R. Shrivastava ( ) · S. P. Singh  
Bennett University, Greater Noida, India  
e-mail: [rajesh.shrivastava@bennett.edu.in](mailto:rajesh.shrivastava@bennett.edu.in)

S. P. Singh  
e-mail: [simarpreet.singh@bennett.edu.in](mailto:simarpreet.singh@bennett.edu.in)

M. K. Hasan  
Faculty of Information Science and Technology, Center for Cyber Security, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia  
e-mail: [hasankamrul@ieee.org](mailto:hasankamrul@ieee.org)

send it to the server. These deployments are also vulnerable to an adversary who wishes to reverse engineer their deployment information, modify control flow by injecting malicious code, tamper with existing code [2] or even clone the IoT device to completely takeover node activity [3].

Control Flow Integrity (CFI) [4] is a well known technology to prevent code tampering. But most of the available CFI solution works on desktop systems or cloud architecture in Intel's x86 environment [5]. Most of the IoT devices use ARM architecture (such as Raspberry Pi). IoT devices with Linux kernel use ASLR to randomize memory areas, e.g., stack and heap. The lack of randomization made ASLR less effective against the real-world attacks [6]. ARM architecture uses security feature No-eXecute (NX) in CPUs to segregate data and code section of memory, which was introduced in ARM v6 and used in all later versions. Despite the available protection methods such as ASLR and NX, adversaries are still able to bypass control flow by using Code Reuse Attacks (CRA), e.g., by using Return Oriented Programming (ROP) [7]. CRA compromise the control flow by using existing vulnerability such as buffer overflow [7], integer overflow, etc. Hence, code tamper-proofing still an open research area to find an effective method to prevent control-flow hijacking in IoT devices. The proposed approach collects information on all vulnerable code points from the binary. We prepare a new execution sequence with the help of ROP gadgets which help to execute code from both data section and code section. The ROP chain (collection of control flow statements and gadgets) executes next instruction and validate itself every time in memory while execution gadgets. If an adversary tries to tamper control flow, then the arrangement of an ROP chain breaks and terminates the execution. Default execution of a program executes from the code section of a binary. The notable contribution in this paper is as follows:

- We use static and dynamic analysis approaches to find vulnerable and control code points such as indirect call/jump instructions, memory leakage points and tainted input.
- Design and develop an ROP based solution for ARM architecture to prevent control-flow tampering and also verify the ROP chain in memory by using user-level probes.

The remainder of this paper is organized as follows. Section 2 reviews the recent work in defense mechanisms against code tampering attacks. Section 3 includes the details of the proposed solution and describes different algorithms used therein. Section 4 includes an evaluation of the proposed solution as well as a theoretical overhead calculation. Finally, Sect. 5 finishes with concluding remarks on the work.

## 2 Related Work

Code reuse attack uploads malicious input into the stack, and it fetches to the code segment through vulnerable functions such as buffer overflow, corrupt pointer, illegal memory access, etc. Code reuse attack broadly categories in two categories ROP

[7] and Jump Oriented Programming (JOP) [6]. To defend against the code reuse attacks in Intel architecture several researchers proposed various solutions [5, 8]. This paper focus on IoT devices having ARM architecture. Many countermeasures to protect a binary program against code reuse attacks have developed during the last decades. CFI [5] is a one available solutions for code tamper-proofing. CFI restricts a program's execution to follow its CFG. The CFI method inject static labels into indirect branches and target instructions, and verifies these label at the time of execution. If the source and destination have different labels, then it treated as a security breach or code tampering. After that, many researchers proposed various solutions for CFI. Walls Robert et al. [9] proposed a security system on ARM Cortex-R devices. In that method, they protect runtime structures, isolating processes, and instrumenting compiled ARM binaries with CFI protection. They used a shadow stack method to prevent CFI. They have to set the size of a shadow stack manually and suffer from a shadow stack overflow problem. That method needs additional space in resource constraints devices. Shrivastava et al. [4] proposed a dynamic canary based method to prevent code tampering. In this method, author inserted dynamic canary values into stack frame these values dynamically generated by the compiler and verifies at the time of execution. But dynamic canary method need manual effort to do the thing and mostly used in Intel's x86 architecture.

### 3 Proposed Approach

#### 3.1 Identification of Vulnerable Code Points

We use following two-step method to identifies these vulnerable code points:

**CFG Analysis** Generates a control flow graph (CFG) of the program binary and collect information of all jump, call, and return addresses. In this method, we divide the whole program into basic-blocks. We used Dynamo Rio to [10] to generates control flow graph. Algorithm explained the execution of the first step.

In this step, we perform static analysis and dynamic analysis in a program binary. The outcome contains an addresses of entry and exit points of basic blocks. But in a program there are huge number of basic blocks available. Securing all the basic-blocks generates enormous overhead, and practically it is not feasible. Hence, Algorithm collect all basic-blocks information that belongs to the control flow graph. Algorithm starts from the main function and traverse the graph through a depth-first algorithm save the path into a hash table. While storing the path we excludes library and system generated address. We only consider those address which present in real program. A path consists of the list of traversed instructions and traversed basic blocks. The Algorithm terminates, if no more predecessor basic block exists, or a function call instruction has been found.

**Information Flow Analysis** The address collected in previous step are huge in number. So in this step we try to prune away non vulnerable code points. For this purpose, we perform information flow analysis over the basic blocks identifies in the first step. In this analysis, we trace each input variable and the basic blocks affected by the entered input values. An adversary uses a code reuse attack by injecting malicious payload into the binary. To inject malicious payload they analyzes source code through taint analyses [7]. Taint analysis helps us to collect all vulnerable points, which helps the adversary to inject malicious payload. We also used same technique to collect all vulnerable code points. After this step, we have all the targeted basic-blocks address which fall within the code regions and vulnerable also [11]. ARM-tracer is used to perform information flow analysis (taint analysis).<sup>1</sup>

---

**Algorithm 1:** Control flow information through basic blocks
 

---

```

Input: Current basic blocks address BB.
Output: Basic Blocks belongs into Control Flow.
isStart = ( BB == startaddress )
if (isCalladdress(BB) or ( !isStart and !libAddress ) ) then
| return false;
end
else
| prevAddress = isStart ? null : getAddress();
| path = prevAddress == null ? new path : NextAddress();
end
if ( path.length() == count(BasicBlocks) ) then
| return false;
end
else
| path.add(current address);
| savePath();
end
return true;
```

---

### 3.2 ROP Chain Augmentation

Once the vulnerable code points identified, next task is to prevent these security sensitive code points. Here, we address the code integrity problem by using an ROP chain augmentation. The ROP based code protraction framework starts with a verification code. The verification code selects from a code fragment at the source or binary level as a start point. We also determine a set of instructions and functions addresses to be protect in a previous step. The proposed method arranges the selected addresses into an ROP chain. This ROP chain uses gadgets addresses to call the next function or instruction. As we already discussed in earlier section, most of the IoT

---

<sup>1</sup> <https://github.com/scotty-kdw/ARM-Tracer>.

devices use ARM architecture (RAspberry pi, Lora, etc.). ARM architecture do not have “ret” instruction. Since ARM uses RISC instruction set where program counter (PC) and link register (LR) are used to hold return address. So, with the help of PC register and LR register we can frame ROP chain. In the ROP chain, after executing a gadget (instruction), the next instruction address is pop to “PC or LR” register. We use these “PC or LR” instructions as our gadget to move from one address location to the next.

In this paper, ROP is a security technique used to execute code in combination with instruction pointer. ROP obtain the control of the call stack and control the flow of existing trusted program binary running on the computer.

Lets discuss proposed solution with an example. Here, Fig. 1 presents disassembled code of a function “fun1”. Fun1 is an ROP initiator function and it is responsible to invoke ROP-chain. Here, the addresses “0x40077474” (add), “0x40077478” (pop), “0x4007747c” (bx), are act as ROP gadgets. These gadgets initialize ROP-chain. The instruction add r0, r4, #6 set the ROP chain address into r0 (i.e., r0 = r4 + 6). After setting the r0 using add instruction, the next instruction is pop instruction. This instruction pop a value into r4 and LR. We are only interested in the value of the “LR” register. Link register (LR) contains the address of the next valid function. The functions address stores in the read-only data section. Once the start string sets accordingly in the “R0” register, then our ROP chain takes control of execution. If an adversary tries to tamper this ROP chain, then the execution addresses are changed, and ROP chain halts. Since the address of gadgets and security-sensitive code points looks the same. It is difficult to identify the control flow of the program. Hence, we achieve our first objective to obfuscate the security framework from adversaries. The ROP initiator function on the stack holds initial control over the control flow and provides the arguments for the first gadget. After the ROP chain initiate, gadgets sequence used in the stack to execute the program and calls the next gadget.

```

0x40077458 <fun1+0>: ldr   r3, [pc,#32]; 0x40077480 <fun1+40>
0x4007745c <fun1+4>: push  {r4, lr}
0x40077460 <fun1+8>: ldr   r4, [pc,#28]; 0x40077484 <fun1+44>
0x40077464 <fun1+12>: add   r3, pc, r3
0x40077468 <fun1+16>: add   r4, r3, r4
0x4007746c <fun1+20>: mov   r1, r4
0x40077470 <fun1+24>: bl    0x40077638 <fun1_r>
0x40077474 <fun1+28>: add   r0, r4, #6
0x40077478 <fun1+32>: pop   {r4, lr}
0x4007747c <fun1+36>: bx    lr
0x40077480 <fun1+40>: muleq pc,r4,r11
0x40077484 <fun1+44>: andeq r3,r0,r8,lsr #4

```

**Fig. 1** Disassemble code with ROP chain

### 3.3 Verification of ROP Chain

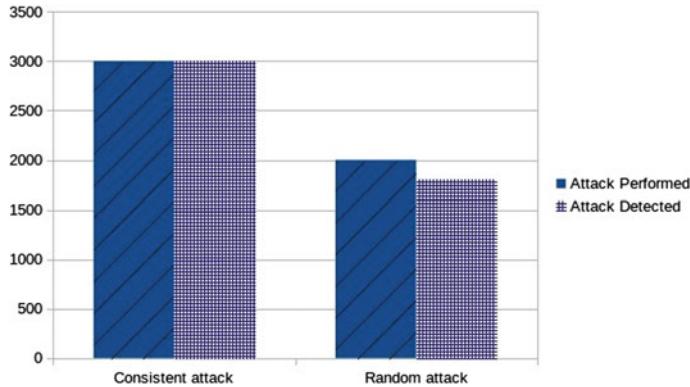
This section helps to verify the ROP chain in memory using user-space probes (U-probe). Continues monitoring helps to prevent ROP chain from any modification. U-probe is a flexible, user-space probing mechanism that creates and manages probe-points user applications. A U-probe user writes a kernel module to specify the desired probe-point of the process and virtual address as a probe, a probe handler to run when the probe-point hits. “U-probe” allows setting probes at different instructions in binaries. Once “U-probes” is enabled, all the target instructions replaced with the tracer instruction (which mostly is “nop”). Once a probe hits, the handler function executes, followed by the U-probe instruction. The “nop” probe tracer stays as it is in memory. This tracer ensures that multi-threaded applications, where two threads can access the same instruction, do not miss a probe hit. The overhead for “U-probe” is much lesser than traditional tracing methods. Also, “U-probe” supports multi-threading and multiple instances of the same binary running. We could differentiate between the different processes that hit the same probe. We attach probes with ROP chain, which hit when the function call by the program. This method helps us to monitor the ROP chain at the time of execution. Hence, we achieve our second objective, i.e., prevent ROP chain from code tampering.

## 4 Result Analysis

### 4.1 Security Evaluation

To evaluate the code tampering detection capability, we have implemented two attack methods: static and random attacks. In the static attack method, we apply our protection method on a vulnerable program and test it with ROP and JOP attack. More specifically, we corrupted the stack pointer by using buffer overflow vulnerability. Since we already protect all vulnerable code points, so change in a gadget identified by “U-probe”. These attacks disturb the order of the gadgets, so it halts the program. Our method correctly detected the existence of such CRA attacks. Similarly, in the random attack, our attack program corrupted random indirect jumps in the binary program. Our method, again, successfully detected such attacks. Figure 2 shows the attack analysis. Static attacks almost identify by 100% accuracy, and the success rate to identifies random attacks is 90%.

To analyze the reliability of the proposed method, we analyzed the false positive and the false-negative rate. For this purpose, we deploy “U-probes”. Uprobes validates all ROP gadgets in memory. All cross-function jumps induced by a function call or a return instruction. Since every legal function call should transfer the program control to entries of the target functions, they must follow the CFI rule. Similarly, all legal returns jump back to the call sites, which must be activated already in the

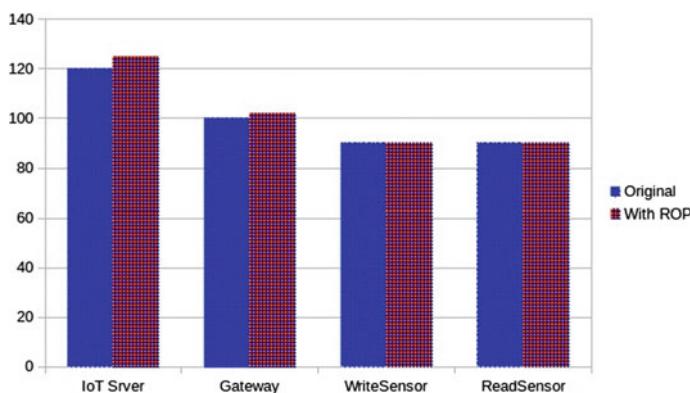


**Fig. 2** Attack evaluation

ROP chain. There are no false positives and does not make a false alarm on regular program executions.

#### 4.2 Performance Overhead

There is no storage overhead due to ROP gadgets because gadgets are already present in the binary. Since gadgets are already part of a binary, hence these gadgets do not generate overhead at the time of execution, the only overhead generated by the verification function, which is negligible. Figure 3 shows the experimental result. We test our security solution with four different programs deployed in different IoT nodes. The four programs are server, gateway, write, and read for IoT devices. The



**Fig. 3** Runtime overhead

result shows that the runtime overhead is negligible. In various execution sequences, we encounter at most 5% overhead. We use the Linux Perf tool to measure overhead.

## 5 Conclusions and Future Scope

In this paper, we proposed a method to prevent code-reuse attacks and a verification technique from defending against code reuse attacks. In this work, we used ROP gadgets to make stack executable and executes all vulnerable functions from the stack through the ROP chain. Our method effectively protects code, but it needs manual efforts. In the future, we come with an automated solution that covers other embedded devices.

## References

1. SANS Institute InfoSec Reading Room, The 2018 sans industrial IoT security survey: shaping IIoT security concerns
2. H. Md. Mahmud, M. Fotouhi, R. Hasan, Towards an analysis of security issues, challenges, and open problems in the internet of things, in *2015 IEEE World Congress on Services (SERVICES)* (IEEE, 2015), pp. 21–28
3. S.P. Singh, A.R. Maini, Spoofing attacks of domain name system internet, in *National Workshop-Cum-Conference on Recent Trends in Mathematics and Computing (RTMC)* (2011)
4. R.K. Shrivastava, K.J. Concessao, C. Hota, Code tamper-proofing using dynamic canaries, in *2019 25th Asia-Pacific Conference on Communications (APCC)* (IEEE, 2019), pp. 238–243
5. S. Rajesh Kumar, C. Hota, Profile-guided code identification and hardening using return oriented programming. *J. Inf. Secur. Appl.* **48**, 102364 (2019)
6. D.L. Vincenzo, A. Dmitrienko, S. Nürnberger, A.-R. Sadeghi, Gadge me if you can: secure and efficient ad-hoc instruction-level randomization for x86 and arm, in *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security* (ACM, 2013), pp. 299–310
7. S. Rajesh, C. Hota, P. Shrivastava, Protection against code exploitation using ROP and check-summing in IoT environment, in *2017 5th International Conference on Information and Communication Technology (ICoICT7)* (IEEE, 2017), pp. 1–6
8. S. Rajesh, C. Hota, G. Mittal, Z. Akhtar, Undertracker: binary hardening through execution flow verification, in *International Conference on Secure Knowledge Management in Artificial Intelligence Era* (Springer, 2019), pp. 3–20
9. J. Walls Robert, N.F. Brown, T. Le Baron, C.A. Shue, H. Okhravi, B.C Ward, Control-flow integrity for real-time embedded systems, in *31st Euromicro Conference on Real-Time Systems (ECRTS 2019)* (Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019)
10. C. Guillermo, C. Gorgovan, M. Luján, Optimising dynamic binary modification across 64-bit arm microarchitectures, in *Proceedings of the 16th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments* (2020), pp 185–197
11. B. Ferdinand, D. Gens, P. Jauernig, A.-R. Sadeghi, E. Stapf, Sanctuary: arming trustzone with user-space enclaves, in *NDSS* (2019)

# Conversion of Punjabi Sign Language Using Animation



Retinder Kaur and Suresh Kaswan

**Abstract** Nature has given the most precious gift to human being is the ability to listen, speak and talk. But there are some unfortunate ones who are unprivileged from this valuable gift. This creates a gap between the normal and underprivileged humans. Sign Language (SL) is the special language that is used by Deaf and Dumb people for communication. SL's is a language which contain symbolic encoded message for communication without speech channel. In this research paper, we have focused on a Punjabi Sign Language. To develop this system we worked on daily used words of Punjabi Language. It consists of components such as Adobe Illustrator, Adobe Flash for generating SL of Punjabi Language. In this we used Illustrator for creating graphics of SL and animate them in Flash. We tested this process and results are coming in positive manners.

**Keywords** Punjabi sign language · Animation · Machine translation

## 1 Introduction

The sign language (SL) is a way of conversation to express views, opinion and thoughts by deaf and dumb person with others in society especially with family members. Mostly people don't know about this language so they are not able to understand it. Many researchers are doing research work on various sign languages of world and in India there is an Indian Sign Language. But during the research on Punjabi Sign Language it has found that there is a few works on it. There is also lot of issues to work on Punjabi Sign Language and one main complication is dictionary of Punjabi Sign Language.

In this period of advanced technology, where graphics and animation become an important part of our day life, so with this technology try to do efforts for making the disables more independent in life and create a bridge between normal, deaf and dumb people. There are many methods adopted by researchers for developing this

---

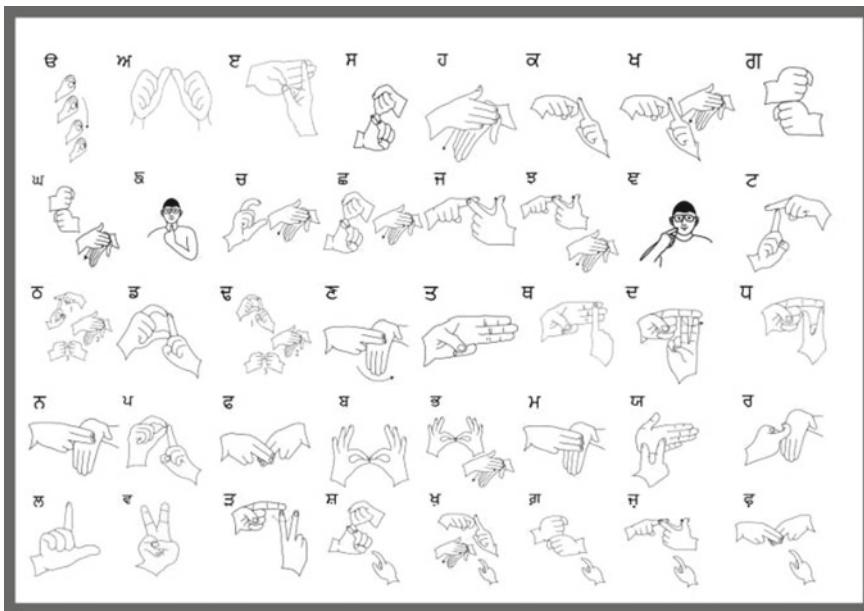
R. Kaur (✉) · S. Kaswan  
RIMT University, Mandi Gobindgarh, Punjab, India  
e-mail: [retinder@gmail.com](mailto:retinder@gmail.com)

language such as gestures of hand, video and real time based technique. Researchers [1–3] have done their research on static images, used gesture for 3D animation. In this paper we worked on Punjabi Sign Language, it is the amalgamation of graphics and 2D animation. In present era, animation plays a vital role and it has impact on each field such as in education, medical, marketing and IT Industry etc.. Also we observe that people of any age show their curiosity in animation work, this is reason that now not only cartoons for children are made in it but also movies such as Avatar, Chaar Sahibzaade and many more made by using this technique. Animation also becomes an easy mode for people to visualize and understand things properly. So we decided to do work on Punjabi Sign Language by using animation. Here we work on 2D Animation. First signs are designed in the form of graphics then these graphics convert into animation.

## 2 Review of Literature

Many types of methods have been used by researchers for developing Sign Language in different languages such as in America, researchers have done research on their sign language named that language as American Sign Language (ASL), same is in British they named the language as British Sign Language (BSL), and in India they named it as Indian Sign Language (ISL) and in Punjab there is Punjabi Sign Language. So according to these, researchers are doing their research and named the sign language based on their country and region's language.

Many researchers [4–6] have worked on video and real time techniques. Chou [4] used technique based on Hidden Markov Model; in this he introduced both hand gestures for sign language. It has observed in various researches that sign language is not same everywhere. Few of sign languages follow single hand technique and in few sign languages used both hands. Both Neural Network's features and Hidden Markov Model are acknowledging hand motions in video. Skin filtering technology [5] based on filter the skin along with ANN. It came to notice that the percentage of success was near about 92%. In this research mainly focused on different gestures. As per geographical area, sign language and spoken languages are used for communication [6]. Dasgupta et al. [7] has done his research on converting text of English language into ISL system. In this he followed Lexical Functional Grammar (LFG) and work on rules of grammar then introduced them into ISL syntax. For this he collected 208 sentences of English and converted them into signs. Verma and Kaur [8] has done their research on Punjabi language, in which they followed HamNoSys with the compilation of near about 235 words and checked them by using the JA SiGML Player app. Dhanjal and Singh [9] has done their research on Punjabi Sign Language. In this they convert the Punjabi words into Indian Sign Language using 3D animation. They worked on making automatic translation machine and the system was designed with the help of ASP.net. M. Alnifai, S. Sampali [10] worked on communication barriers and on communication apps those are working for deaf and hearing impaired people. During the research it has found that near about 6 apps out of 55 communication



**Fig. 1** Varnmala of Punjabi words along with sign language

apps are properly designed for those people. According to their research it becomes necessary to develop more apps that helps to attain the needs of impaired people and reduce the communication barriers.

### **3 Research Methodology**

### **3.1 Punjabi Sign Language**

Punjabi Sign Language has developed for deaf and dumb people so that they can interact with normal person. This language based on single and both hands for communication. In this paper focused on 35 alphabets (consonants and vowels) (Fig.1), numbers (1-10) and some daily used words.

### 3.2 *Methodology*

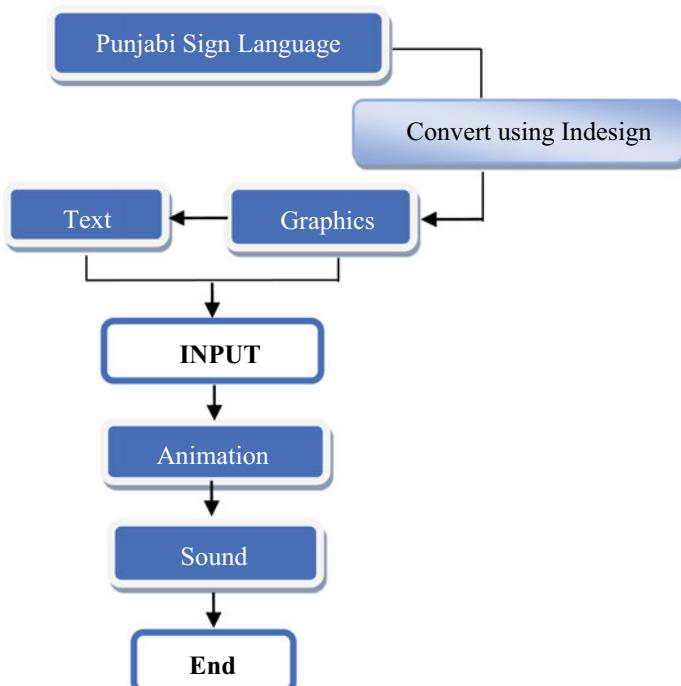
Researchers from all over the world, has done a lot of work on Sign language. Using Automatic machine translation is a complex task because there are basically

two steps: the translation component, and recognition of sign language or synthesis gesture that depends on the direction of translation process.

In research first we convert the Punjabi Sign Language into graphics then these graphics convert into animation. We used PAL standard for Animation of sign language. PAL is Phase Alternative line which is used in Asia for creating animation. PAL works on 24 fps. “Fps” means frames per second so 24 frames per second are used to create animation. Along with we used frames, keyframes and inbetweens to create animation. Sound of every sign of Punjabi language also put in it during animation.

### 3.3 Output Module

This following diagram explains the method to convert the sign language into animation. This all process is based on 2D. Here we used 2D animation. We used softwares those are related to graphics (Adobe Illustrator) and Animation (Adobe Flash) for this research (Fig. 2).

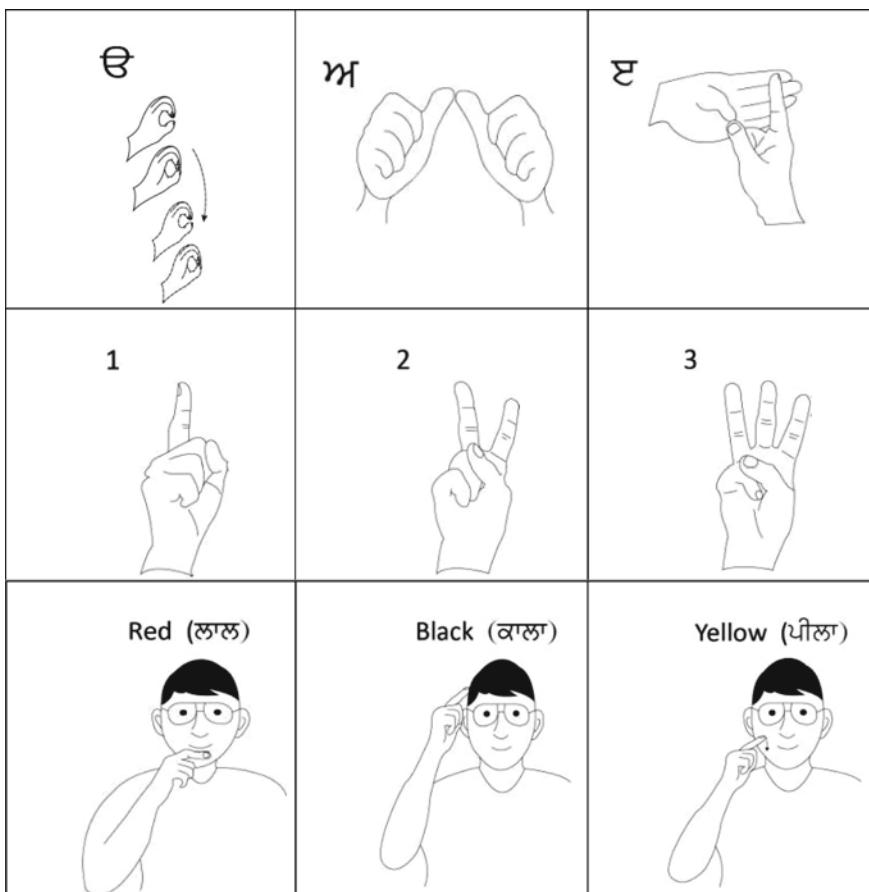


**Fig. 2** Flow chart of conversion of sign language to animation

## 4 Implementation

To start the work on it, first we design the graphics of Signs. These designs are created in Adobe Illustrator. Adobe Illustrator is one of the software to create designs. After designing the signs, next step is to take output of these in the form of .jpeg, .png, etc. These signs are in form of still images and we can export illustrator files in Adobe Flash (Fig. 3).

Second step is to convert the graphics into animation. So for this, import still images of signs or directly export adobe illustrator file in Flash software. Adobe Flash is 2D based animation software. Here we work on animation to animate the signs. For animation we use keyframes, inbetweens for smooth animation also use tweens such as motion tween. Tweens helps to create animation easily. We follow two methods to create animation that is tweening and step by step animation.

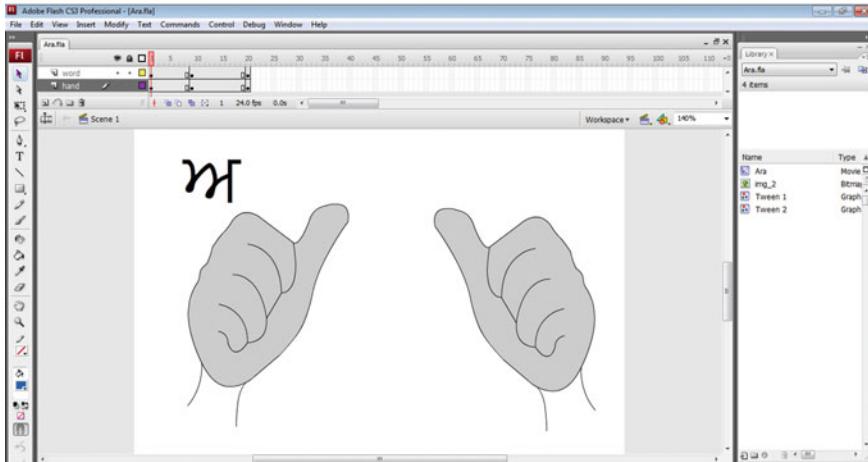


**Fig. 3** Graphics of Punjabi sign language

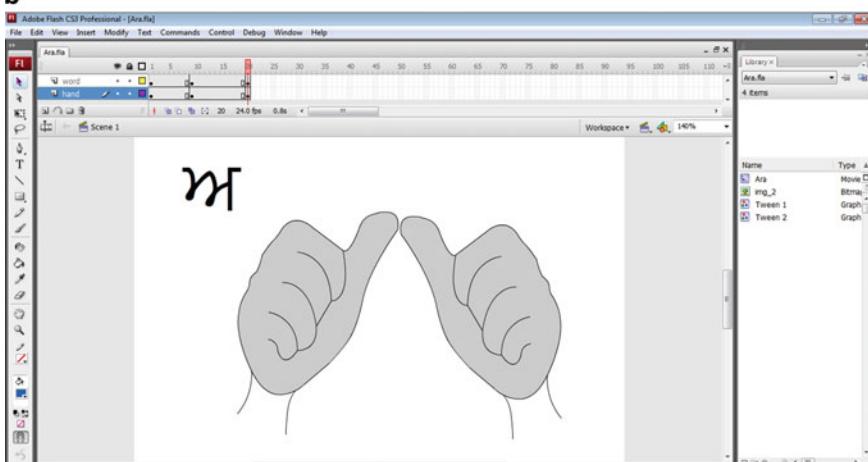
Tweening is applied on main frames (keyframes). But step by step animation can do on each and every frame. On each frame we can move the character or character's arms, hands, work on expressions. So according to demand we use both methods by following the principles of animation. Principles of Animation have 12 rules those we follow to create animation both in 2D and 3D. Without these rules we can't imagine animation. These are like guidelines which help to animate the object.

These Figs. 4, 5 and 6 we try to show that how to create Signs and animate them by using keyframes, motion tween and step by step animation. Here in each frame

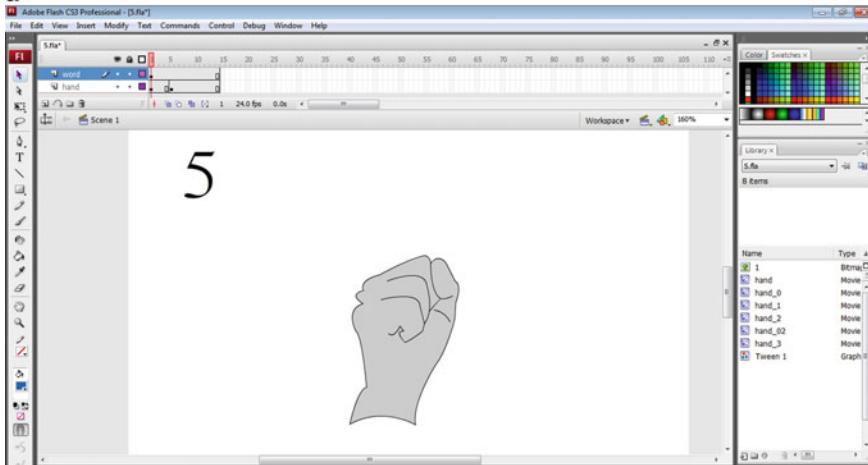
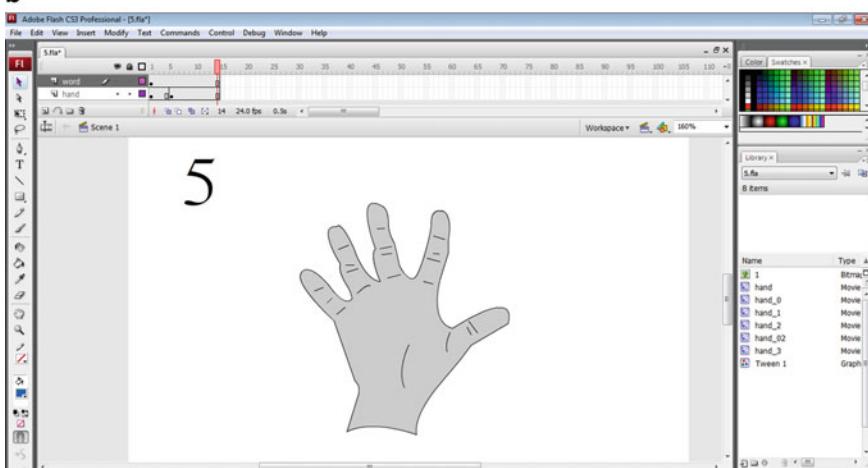
a



b

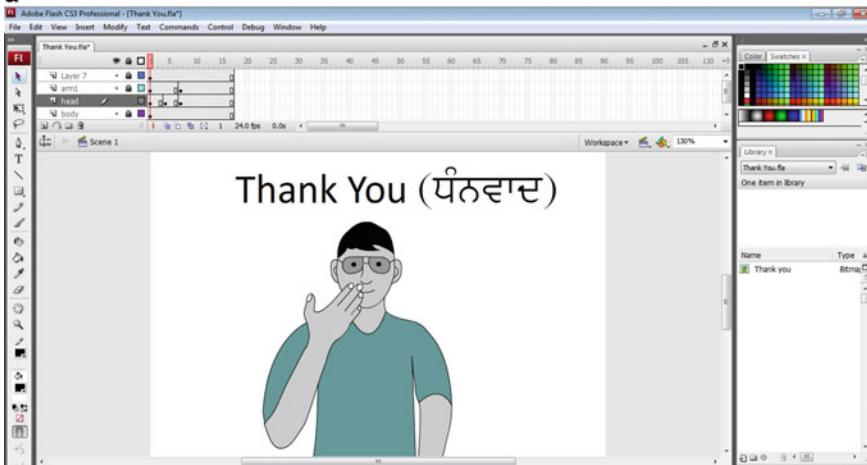
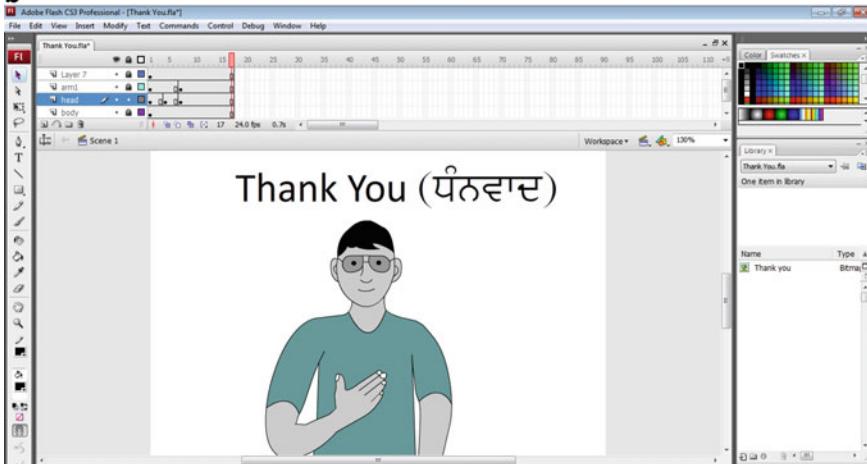


**Fig. 4** In these figures show how to work on frames for animation. First step hands are on first main frame, in second step hands are on next main frame. So when we play them on 24 fps then it creates motion between these hands those are showing in these two figures

**a****b**

**Fig. 5** In following figures, a numerical number “5” in which we see that on first frame hand looks like this and in next frame hand is open so when we animate them on PAL standard, it creates motion

we create new pose or move of character’s arms, hand etc. then animate them by following PAL standard.

**a****b**

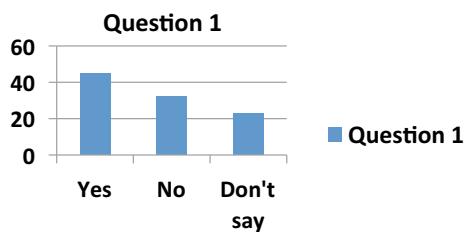
**Fig. 6** In following figure, in first frame hand is near to mouth then in next keyframe hand is away from mouth and in front of body. When we animate them it creates illusion of motion

## 5 Results and Discussion

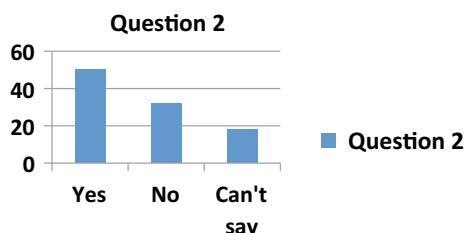
With the main objectives keep in mind, the researcher has followed a questionnaire. The researcher has studied the content and characteristics of 2D animation on both deaf and dumb and normal people. Thus the questionnaire has designed to keep the main objectives in mind. We have tested on 25 people based on Punjabi language alphabets, numbers and some daily used words. So we create some questions related to research in which our first based on (Figs. 7, 8, 9, 10 and 11):

1. Do you think to convert Sign Language to Animation is beneficial? (Fig. 7)

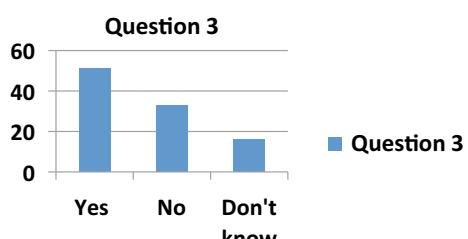
**Fig. 7** Shows that 45% of the people think that it is beneficial to convert sign language by using animation



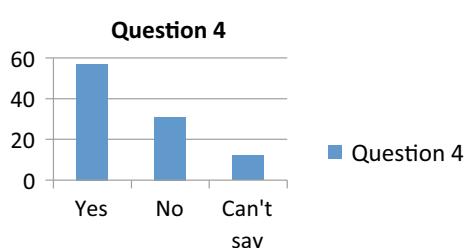
**Fig. 8** Shows that 50% people say it creates a bridge between normal, deaf and dumb people



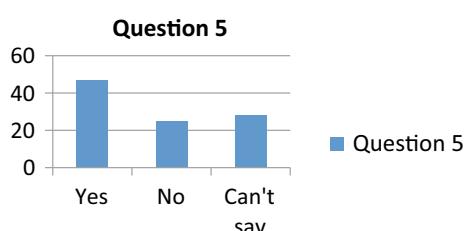
**Fig. 9** Shows that 51% people think that animation helps to make sign language easy for normal person



**Fig. 10** Shows that 57% people think that it helps to understand this language easily by abnormal people/children in an initial stage



**Fig. 11** Indicates that 47% people think that animation helps to increase learning capability about sign language



2. Will conversion of Sign Language bridge a gap between normal, deaf and dumb people? (Fig. 8)
3. Will Animation be an easy way to understand Punjabi Sign Language by normal people? (Fig. 9)
4. Is conversion of sign language using Animation helps deaf and dumb people/children to understand this language easily in initial stage? (Fig. 10)
5. Do you think Animation helps to increase learning capability of people about Punjabi Sign Language? (Fig. 11)

## 6 Conclusion and Future Scope

As we come to know after study that this technology is about convert the sign language using animation so that every person can able to learn this language easily. In research we work on App for conversion of Punjabi Sign Language. In this app, with the help of graphics and cartoon character we are trying to show animation of Sign language of Punjabi word along with its voice. So that each and every one recognize the word properly. Animation helps to make sign language easy for everyone; Secondly children are very fond of animation related work such as cartoons, animated movies. Children always tried to copy them and easily grasp cartoons activities. Third animation is very good in visualization to make tough thing easy for understanding.

Now a days not even children but also elders like animated movies etc. example is Avatar movie. So there is a great scope in this field to make Punjabi Sign language familiar and easy for people. Fourth for teachers those are teaching to deaf and dumb people, helps them to make teaching more interactive, interesting and take less time by students to understand the things properly. It has also a great future scope because technology is developing day by day. Animation becomes a first choice in every field such as Education, Medical, Entertainment, Engineering, Architecture, gaming and in scientific visualization etc. With the help of animation we can make Punjabi sign language easy by creating short movies, apps, games etc. So animation has vast area to explore.

## References

1. M.K. Bhuyan, M.K. Kar, D.R. Neog, Hand pose identification from monocular image for sign language recognition, in *Proceedings of IEEE International Conference on Signal and Image Processing Applications (ICSIPA 2011)*, Malaysia (IEEE, 2011), pp. 378–383
2. J. Singha, K. Das, Indian sign language recognition using Eigen value weighted Euclidean distance based classification technique. *Int. J. Adv. Comput. Sci. Appl.* **4**(2), 188–195 (2013)
3. A.S. Ghotkar, R. Khatal, S. Khupase, S. Asati, M. Hadap, Hand gesture recognition for Indian sign language, in *International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India (IEEE, 2012)

4. F.H. Chou, Y.C. Su, An encoding and identification approach for the static sign language recognition, in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (2012), pp. 885–889
5. M.P. Paulraj, S. Yaacob, S.Z. Azalan, R. Palaniappan, A phoneme based sign language recognition system using skin color segmentation, in *6th International Colloquium on Signal Processing & Its Applications (CSPA)* (2010), pp. 86–90
6. N.U.N. Soogund, M.H. Joseph, SignAR: a sign language translator application with augmented reality using text and image recognition. *J. Appl. Technol. Innov.* 33–37 (2019)
7. T. Dasgupta, S. Dadpat, A. Basu, Prototype machine translation system from text to Indian sign language, in *Proceedings of the IJCNLP-08 Workshop on NLP for Less Privileged Languages*, IJCNLP, Hyderabad, India (2008), pp. 19–26
8. A. Verma, S. Kaur, Indian sign language animation generation system for Gurumukhi script. *Int. J. Comput. Sci. Technol.* 6(3), 117121 (2015)
9. A.S. Dhanjal, W. Singh, An automatic conversion of Punjabi text to Indian sign language. *EAI Endorsed Trans. Scalable Inf. Syst.* 7(28), e9 (2020)
10. M. Alnfiai, S. Sampali, Social and communication apps for the deaf and hearing impaired, in *2017 International Conference on Computer and Applications (ICCA)* (IEEE, 2017), pp. 120–126

# Deep Learning Based Classification of Microscopic Fungal Images



Amit Sharma, Ajay Lakhnotra, Jatinder Manhas, and Devanand Padha

**Abstract** Early prognosis of fungal contagium may be based on minute examination using microscope. In most cases, however, it becomes unfit for the abstract identification of species because of their apparent similarity. So, it becomes absolute necessary to employ more biochemical tests. In order to detect and identify the nine fungal species from the microscopic images, transfer learning has been deployed by the authors without data augmentation and got 95.45% classification accuracy. Data augmentation has also been applied on the dataset under consideration then fed into the network and got 94.77% classification accuracy. Inception V3 network was used for the study.

**Keywords** Fungi · Inception net · Deep learning · Machine learning · Transfer learning · CNN

## 1 Introduction

Fungi are the third largest kingdom of eukaryotes that digests food externally and absorbs nutrients directly through its cell walls. Fungi are omnipresent. There is exactly no data about the number of species of fungi. According to one estimation, their number is between 2.2 and 3.8 million. Till date only 120,000 have been documented. Fungi are considered as plants as they grow in soil and do not move like animals. The common examples of fungi are white mushrooms and shiitake. Fungi

---

A. Sharma (✉)

Department of Computer Science and I.T, Kathua Campus, University of Jammu, Jammu and Kashmir, India

e-mail: [amitsharma@jammuuniversity.ac.in](mailto:amitsharma@jammuuniversity.ac.in)

A. Lakhnotra · J. Manhas

Department of Computer Science and I.T, Bhaderwah Campus, University of Jammu, Jammu and Kashmir, India

D. Padha

Department of Computer Science and I.T, Central University of Jammu, Jammu and Kashmir, India

are not capable of carrying out photosynthesis by themselves. They obtain their carbon and energy from other organisms [1]. The fungi that obtain their nutrients from plants or animals are called biotrophs and others that obtain nutrients from dead plants or animals are called saprotrophs [2]. Some fungi kill host cells in order to obtain their nutrients; these are called necrotrophs [3]. Fungi has so many practical uses like in fermentation, creation of modern medicines, dyes, food items, folk medicines, pest controls.

Fungi are both useful and harmful for human beings. On the positive side, they are used to create a wide range of industrial products. They are also used to create many antibodies, statins [4] and anti-cancer drugs [5]. On the negative side, fungi have many harmful impacts on humans. According to one survey, around 25% of the world's population suffers from a fungal infection of the hair, skin or nails each year, such as athlete's foot. Fungal infections kill three times more people than malaria. Apart from this, the matter of great concern is the infections caused by the fungi that are normally harmless. The people with weak immune system are the victims of these fungi. The latest statistics suggest that at least 1.6m people die every year as a result of the immunosuppressive drugs [6]. Fungi also cause damage to the crops, timber and food. The moulds of fungi like Penicillium and Aspergillus, spoil many stored foods. We taken up our study to detect and recognize nine species of fungi namely *Candida-albicans*, *Candida-glabrata*, *Candida-lusitaniae*, *Cryptococcus-neoformans*, *Candida-parapsilosis*, *Candida-tropicalis*, *Malassezia-furfur*, *Saccharomyces-boulardii* and *Saccharomyces-cerevisiae*.

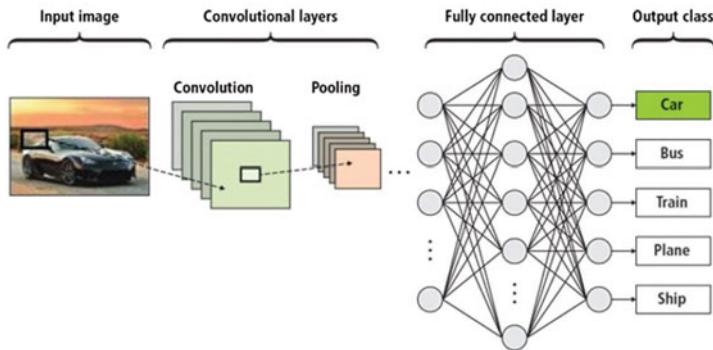
## 2 Related Work

Zielinski et al. [7] came up with a mycological diagnosis of fungi from microscopic images. A collection of 180 microscopic images DIFaS of nine species of fungi were taken for performing the experiment. More than 14 variants of already tuned deep learning models like AlexNet, DenseNet169, InceptionV3 and ResNet were used and got more than 82% recognition accuracy. A smart classification framework with transfer learning for recognition of nine species of *Aspergillus* genus was proposed by Mital et al. [8]. In the proposed framework 1152 images with nine classes were specified as input. Different deep learning models were analyzed experimentally. It was found that MobileNet V2 got the highest categorization accuracy of 96.66%. Billones et al. [9] deployed a deep learning model to detect the different fungal species from macroscopic images of *Aspergillus* fungi. The authors used the dataset consisting of 4545 images containing 9 types of fungus. After data augmentation and data preprocessing a simple convolutional neural network was deployed to classify the species with the classification accuracy of 90.06%. An intuitive diagnosis system to recognize the fungal keratitis in the human eye using microscopic images was developed by Lv et al. [10]. A total of 2088 images were used in the study which contains 688 positive examples and 1400 images have negative examples. A deep residual network (ResNet) with almost 101 convolutional layers was used by

the authors for diagnosis with an accuracy of 93.64%. Billones et al. [11] gave an intelligent framework based on deep learning to differentiate the fungi species from microscopic images of Aspergillus. A dataset having 4545 images containing nine species of fungi was exercised in the study. Data augmentation and data preprocessing was done and then a convolutional neural network was deployed to recognize the species with the accuracy of 87.50%. Development of an automated clinical detection system for the detection of microbial keratitis from corneal pictures was done by Kuo et al. [12]. 228 pictures of cornea were taken in which 114 were positive and 174 were negative samples. A convolutional neural network was trained on the dataset and has given 70% classification accuracy. Dawood et al. [13] has developed a diagnostic model for the detection of eatable and non edible mushrooms. 22 parameters were supplied as an input to the multilayer perceptron with back-propagation having 8124 records. The configuration of the 3-layer Network having 22 input neurons, 2 hidden neurons and 1 output neuron. Experimental results asserted that the classification accuracy was close to 100%. A novel dataset consisting of 2414 digital images of wheat fungi diseases has been created to predict the disease by Genaev et al. [14]. An Efficient Net architecture based on CNN was put up for the disease detection caused by fungi on wheat plants and was implemented on the WFD2020 with more than 94% prediction accuracy. An highly unbalanced, fine grained and benchmarked dataset, Danish Fungi 2020 (DF20) consisting of 295,938 images containing 1604 classes was developed by Picek et al. [15]. DF20 contains the images gathered from 30 countries consisting of Fungi, Protozoa and Chromista kingdoms. CNN and Vision Transformers were deployed in the categorization of the above fungi species in the study. In this study CNN based SE-ResNeXt-101-32X4d were compared with Vision Transformer architectures Vit-Large/16 and ViT-Base/16 and it was found that Vision Transformer architectures had achieved 81.25% classification accuracy. Ma et al. [16] came up with an automated system based on deep learning to differentiate the seven fungi species of Aspergillus from microscopic images. A dataset consisting of 8995 images containing seven species was used in this study. A total of 17142 images created from data augmentation and data preprocessing. Xception model based on convolutional neural networks was deployed to identify the species with the accuracy of 99%. An intelligent diagnostic model was created for recognition of disease caused by fungal infection on leaves of apple plants by Gaikwad et al. [17]. 9164 images were taken for the four classes The CNN model was trained on the dataset and authors achieved the accuracy of 88.9%.

### 3 Deep Learning

It is a matter of great astonishment how Google translates the whole paragraph written in one language to another in fraction of seconds. How YouTube and Netflix recognize our taste in movies or videos and provide us with perfect suggestions? How can cars drive by themselves? Well, these all are the outcomes of deep learning. Deep learning is a sub part of machine learning which further a sub part of artificial intelligence is.



**Fig. 1** Architecture of deep learning

AI is a very broad term and it refers to the approach that lets computers to imitate human behavior. Whereas machine learning refers to a set of procedures trained on data that makes all of this possible and give the results based on series of inputs. The algorithms in deep learning attempts to draw identical results as a human would constantly examining data with a given logical layout. To attain this, deep learning applies multilayered layout of procedures called neural networks. Neural network is a different concept altogether. Neural networks are just like our brain. Neural networks can be trained to accomplish the same task on data just like our brain recognizes the patterns and classifies the different type of information. Whenever new information is acquired, the brain compares it with the already familiar objects. Deep neural networks applies the same concept to recognize the object and generate new information [18]. An Architecture of Deep learning is shown in Fig. 1.

### 3.1 Convolutional Neural Networks

Convolutional neural network is a type of artificial neural network that is most widely used for image analysis. CNN can also be applied for other data analysis or pattern recognition problems. CNN is specialized to detect patterns and make sense out of them. This ability of CNN makes it appropriate for analysis of images. It comprises hidden layers called convolutional layers followed by fully connected networks. CNN also has non-convolutional layers but the basis of CNN is the convolutional layers. Like other layers, the input is given to a convolutional layer that transforms it in a way that becomes input to the following layer. The transformation done by convolutional layers is convolution operation [19].

### ***3.2 Transfer Learning***

Transfer learning is a mechanism where a framework designed for a job is reused as the starting point for a model on a second job. Transfer learning is a very famous technique that can train deep neural networks with comparatively less data. Transfer learning applies the knowledge of pretrained machine learning framework to a problem which is distinct but related to the original problem. Transfer learning assists to apply what has been learned in one task to improve the abstraction in another. We impel the weights that a network has obtained at “task A” to a new “task B.” The basic concept is to apply the expertise of a model acquired from a job with existing labeled training data in a new job for which ample amount of data is not available. Rather than initiating the learning process from beginning, we start with results obtained from solving a similar relevant task. The problems in the area of computer vision and natural language processing tasks like sentiment analysis which requires the vast computational power employs transfer learning. Transfer learning is actually not a machine learning technique. It can be considered as a “design methodology” adopted for active learning [20].

### ***3.3 Inception Network***

The Inception network was a crucial landmark in the progress of CNN classifiers. Previous to the inception network, most leading CNNs just piled up convolution layers deeper and deeper, wishing to get finer pursuance. The Inception network on the other hand, was very intricate. It applied a lot of practices to excel the performance i.e. both speed and accuracy. Inception network evolved constantly and led to the arise of various new versions [21]. The popular versions are Inception v1, Inception v2 and Inception v3. Each version is a repetitive development over the previous one. Realization of the progress can root for us to build custom classifiers that are refined both in speed and accuracy.

### ***3.4 Inception V3***

The contribution of auxiliary classifiers is not so much near the completion of the training process, when accuracy was closer to saturation. Inception Net V3 incorporated various upgrades required for Inception V2, and in addition used the Factorized  $7 \times 7$  convolutions, RMSProp Optimizer and BatchNorm for the auxiliary-classifiers [22].

## 4 Dataset

The dataset consists of 180 microscopic images containing nine species having twenty images per class and called DIFaS database. It has 9 strains with 2 preparations and 10 images for each species. The dataset can be reached and downloaded at [https://figshare.com/articles/figure/Sample\\_images\\_from\\_DIFaS\\_database\\_12589923/1](https://figshare.com/articles/figure/Sample_images_from_DIFaS_database_12589923/1), [https://drive.google.com/file/d/1-CjXE-HO3xz36x\\_XNJVDTRV89xDnjLlu/view?usp=sharing](https://drive.google.com/file/d/1-CjXE-HO3xz36x_XNJVDTRV89xDnjLlu/view?usp=sharing) [7]. There were a limited number of image samples in the dataset for each species, therefore, in order to generate a comparatively large number of images for training, testing and validation, a data preprocessing technique called data augmentation was used. The image data generator used various methods like rotation, scaling, shearing, height and width shifting to generate new images from the existing images in the given dataset. After data augmentation a collection of 900 images were generated consisting of 100 images per fungal species.

## 5 Methodology

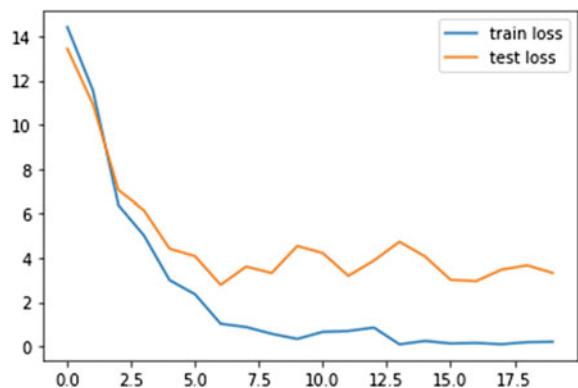
The whole study in this paper was divided into two parts. In the first experiment 180 images without data augmentation were fed into the InceptionV3 network which is already pretrained. In the second part of the study 900 images were given as input to the Inception V3 network. Google Colaboratory with GPU was used as a platform for experimentation. Training and testing was done with 70% and 30% images respectively in both the experiments.

## 6 Results

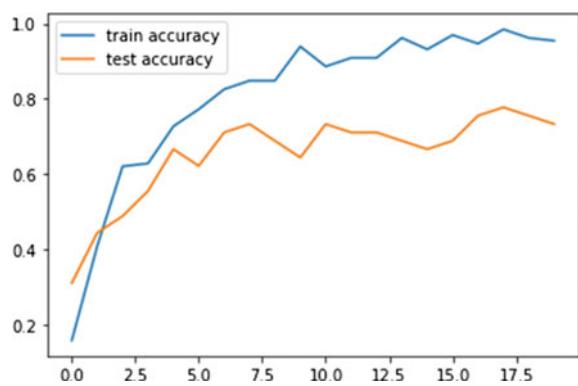
In order to detect and identify the nine fungal species from the microscopic images, transfer learning has been employed by the authors without data augmentation and got 95.45% classification accuracy. Data augmentation has also been applied on the dataset under consideration then fed into the network and got 94.77% classification accuracy.

Figure 2 and 4 shows the loss graphically for the Inception V3 Network during the training and testing phase. Figure 3 and 5 displays the accuracy graphically for the Inception V3 Network during the training and testing phase.

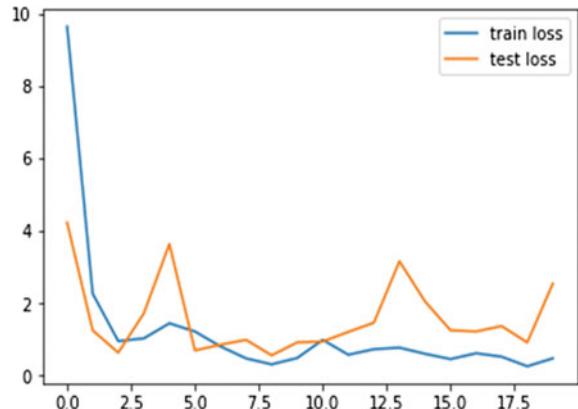
**Fig. 2** Model loss without augmentation



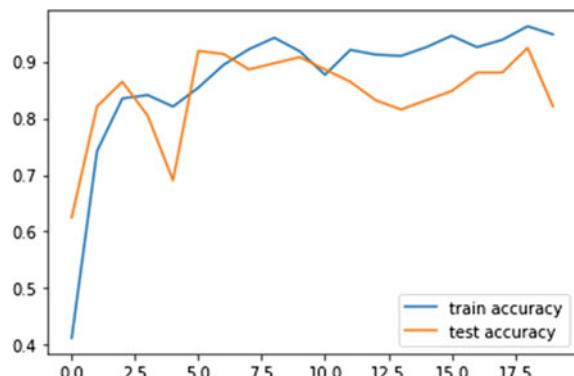
**Fig. 3** Model accuracy without augmentation



**Fig. 4** Model loss with augmentation



**Fig. 5** Model accuracy with augmentation



## 7 Future Scope

In this paper, we applied transfer learning using deep networks for the classification of nine species of fungi. The dataset we used had very limited number of images. In future, the accuracy of the model may be increased by taking the large and more diversified dataset. Also, the number of species of fungi can be increased. Moreover, the data we used was a balanced data whereas in actual scenario, the data is not that much balanced.

## References

1. The Conversation Homepage, <https://theconversation.com/five-facts-that-show-we-know-too-little-about-fungi-and-their-harmful-effects-on-human-health-95741>, last accessed 2021/09/20
2. APS Homepage, <https://www.apsnet.org/edcenter/disandpath/fungalasco/intro/Pages/IntroFungi.aspx>, last accessed 2021/09/21
3. APS Homepage, <https://www.apsnet.org/edcenter/resources/illglossary/Pages/N-R.aspx#necrotroph>, last accessed 2021/09/20
4. Statin Homepage, <https://en.wikipedia.org/wiki/Statin>, last accessed 2021/09/23
5. Anti-Cancer drugs Homepage, [https://en.wikipedia.org/wiki/Anti-cancer\\_drugs](https://en.wikipedia.org/wiki/Anti-cancer_drugs), last accessed 2021/09/24
6. Gaffi Homepage, <https://www.gaffi.org/why/fungal-disease-frequency>, last accessed 2021/09/25
7. B. Zielinski, A. Sroka-Oleksiak, D. Rymarczyk, A. Piekarczyk, M. Brzydacz Włoch, Deep learning approach to describe and classify fungi microscopic images. *PLoS one* **15**(6), e0234806 (2020)
8. M.E. Mital, R.R. Tobias, H. Villaruel, J.M. Maningo, R.K. Billones, R.R. Vicerra, A. Bandala, E. Dadios, Transfer learning approach for the classification of conidial fungi (genus aspergillus) thru pre-trained deep learning models, in *2020 IEEE Region 10 Conference (TENCON)* (IEEE, 2020), pp. 1069–1074
9. R.K.C. Billones, E.J. Calilung, E.P. Dadios, N. Santiago, Image-based macroscopic classification of aspergillus fungi species using convolutional neural networks, in *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (IEEE, 2020), pp. 1–4

10. J. Lv, K. Zhang, Q. Chen, Q. Chen, W. Huang, L. Cui, M. Li, J. Li, L. Chen, C. Shen et al., Deep learning-based automated diagnosis of fungal keratitis with in vivo confocal microscopy images. *Ann. Transl. Med.* **8**(11) (2020)
11. R.K.C. Billones, E.J. Calilung, E.P. Dadios, N. Santiago, Aspergillus species fungi identification using microscopic scale images, in *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (IEEE, 2020), pp. 1–5
12. M.-T. Kuo, B.W.-Y. Hsu, Y.-K. Yin, P.-C. Fang, H.-Y. Lai, A. Chen, M.-S. Yu, V.S. Tseng, A deep learning approach in diagnosing fungal keratitis based on corneal photographs. *Sci. Rep.* **10**(1), 1–8 (2020)
13. K.J. Dawood, M.H. Zaqout, R.M. Salem, S.S. Abu-Naser, Artificial neural network for mushroom prediction. *Int. J. Acad. Inf. Syst. Res. (IJAISR)* **4**(10) (2020)
14. M. Genaev, E. Skolotneva, E. Gulyaeva, E. Orlova, N. Bechtold, D. Afonnikov, Image-based wheat fungi diseases identification by deep learning (2021)
15. L. Picek, M. Šulc, J. Matas, J. Heilmann-Clausen, T.S. Jeppesen, T. Læssøe, T. Frøslev, Danish fungi 2020—not just another image recognition dataset, arXiv preprint [arXiv:2103.10107](https://arxiv.org/abs/2103.10107) (2021)
16. H. Ma, J. Yang, X. Chen, X. Jiang, Y. Su, S. Qiao, G. Zhong, Deep convolutional neural network: a novel approach for the detection of aspergillus fungivia stereomicroscopy. *J. Microbiol.* **59**(6), 563–572 (2021)
17. S.S. Gaikwad et al., Fungi classification using convolution neural network. *Turk. J. Comput. Math. Educ. (TURCOMAT)* **12**(10), 4563–4569 (2021)
18. Machine Learning Mastery Homepage, <https://machinelearningmastery.com/what-is-deep-learning/>, last accessed 2021/10/03
19. Convolutional Neural Network Homepage, [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network](https://en.wikipedia.org/wiki/Convolutional_neural_network), last accessed 2021/10/04
20. Machine Learning Mastery Homepage, <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>, last accessed 2021/10/05
21. Towards Data Science Homepage, <https://towardsdatascience.com/a-simple-guide-to-the-variations-of-the-inception-network-7fc52b863202>, last accessed 2021/10/06
22. GeeksforGeeks Homepage, <https://www.geeksforgeeks.org/inception-v2-and-v3-inception-network-versions/>, last accessed 2021/10/06

# Application of Artificial Neural Network and Multiple Linear Regression in Predicting Consumer Attitude Towards Use of Mobile Wallet



Anju Singh and Ruchi Jain

**Abstract** The persistent expansion of e-commerce and financial technology in last few decades has accelerated the use of mobile technology and innovative payment methods which has further transformed the consumer, commerce and trade. The adoption of mobile wallets had led to the exploration of the factors which affect the attitude of the consumer in adopting the mobile wallets in their everyday life for financial transactions. The study aims to examine the role of various factors in predicting consumer attitude towards use of mobile wallets. The data has been collected from 501 mobile wallet users through structured questionnaire. The proposed research model was first examined using Multiple Linear Regression. The significant results were further examined using Artificial Neural Network. The outcomes reveal that perceived ease of use, security and usefulness affects the attitude of consumer towards use. Moreover, perceived security is identified to be having strong impact on attitude of the consumer to use mobile wallets. The artificial neural networks (ANN) work on machine learning algorithms and mathematical models, which is used on nonlinear complex data to imitate the information processing of a human brain. On the basis of the results it can be concluded that Artificial Neural Network outperforms multiple linear regression based on the prediction of the consumer attitude towards use of mobile wallets.

**Keywords** Artificial neural network · Mobile wallets adoption · Attitude towards use · Perceived usefulness · Perceived ease of use · Perceived security

## 1 Introduction

Financial Technology has commanded a new line of consumer experience in terms of better speed, convenience, flexibility in payments and seamless payments making it a boon for the e-commerce industry. Recent studies reveal that after pandemic the customers approach towards the use mobile wallets has changed aiming towards the

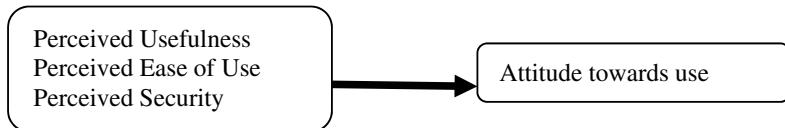
---

A. Singh (✉) · R. Jain  
IIS (Deemed to be) University, Jaipur, India  
e-mail: [anjusingh.choudhary@iisuniv.ac.in](mailto:anjusingh.choudhary@iisuniv.ac.in)

bright future of the mobile wallets. It has been concluded from the recent studies that by 2025 every second person would be using mobile wallets for digital payments. Digital wallets have proved their success in the most difficult times for making any contactless payment and therefore are no more just a trend but have become a habit. This leads to the purpose of the present study which examines primarily the impact of the independent variables adopted from the models of Technology Acceptance through review of literate to measure the impact of perceived ease of use, perceived usefulness and perceived security (external factor) on consumer attitude towards use of mobile wallets. Many researchers have developed a set a framework to investigate the adoption to the technology by the consumer using different technical and scientific algorithms. All these models have been efficient enough to predict and evaluate the adoption of technology and consumer attitude, behavior intention to use technology. Technology Acceptance Model by Davis [1] is one amongst the few popular models. The model measures the factors which are statistically significant in influencing the end user to adopt a technology. The validity of this model has been widely examined in the field of IOT, food technology, agriculture, healthcare, digital banking etc. Further, this study commences in developing a research model based on series of evaluation through multiple Linear Regression and artificial neural network to compare which technique is robust in evaluating the relationship [2]. Neural network is a widely acceptable machine learning technique which used for predicting and works on examining large number of varied inputs and computes the output. It has been effective enough in predicting a variety of problem. This model is a mathematical program which operates similarly to a human brain. Studies conducted using Neural Network has proved its validity in the sub-streams of finance and economics. Financial technology is one such stream where this model can bring together the objectivity of algorithm and psychometric approach to understand the human tendencies in framing consumer attitude and perception.

## **1.1 Research Design**

The present study is exploratory in nature and examines the factors affecting mobile wallet usage. Using a convenient sampling, sample of 501 mobile wallet users in Jaipur City has been drawn for the primary data collection. The self-administered questionnaire has been developed in which information on the demographic profile of the respondents and factors including perceived ease of use, usefulness, trust, security, etc. using Likert scale have been obtained. Two different types of mathematical model (multiple regression analysis model and neural network model) have been constructed using the results of the estimation trials to predict consumer attitude towards the use of mobile wallets. The factor analysis has been used to draw factors for the study. The Cronbach Alpha reliability for each factor was found including, perceived ease of use (0.862), usefulness (0.732), security (0.811), and consumer attitude towards use (0.717).



## 1.2 *Review of Literature*

Many studies have been conducted in the field of finance and related aspects of accounting where the predictions of the ANN and the multiple regressions have been compared. Studies have compared multiple regression analysis with ANN approach for the predictability of the financial strength of banks in India based on financial ratios. The study concluded that ANN model is found to be better in forecasting the CRAR of banks [3]. ANN has also been used in the studies related to the financial technology. Studies using multilayer perceptron technique examined the intention of the consumer to use the peer to peer payment system used logistic regression and proposed an ANN model prediction on the same [4]. Through intensive study on the mobile wallet adoption model concluded about ANN and SEM as an effective hybrid approach in predicting the mobile wallet adoption [5]. The results of ANN model confirm the results of SEM analysis with some variation. Researches have also been conducted for the analyses of the customer chum prediction using three different hierarchical models. The study uses a systematic approach to classify the data churners and non-churners using ANN and regression into [6]. ANN and Multiple Regression has also been used in predicting the performance of banks using ROA as independent variable and seven independent variables and concludes that ANN is a much better tool in the prediction of bank performance [7]. Determinants of capital structure have been analyzed in industries of Taiwan using ANN and multiple regression. The conclusion are drawn on the basis of the RMSE value confirming that ANN model achieves better bit and is able to make better forecast for debt ratios in comparison to regression models [8].

## 2 Discussion and Results

In the present study the measurement model first examine the impact of independent variables (PEOU, PU, and PS) on dependent variable (ATU) through multiple regression analysis. The first step in the process is to clean the data for any outliers. Using Mahalanobis distance method the data was cleaned for outliers. The assumptions of the multiple Regression model were further examined. PP Plot of the standardized residual confirms that the residual are normally distributed. The independent variables found insignificant in the model like perceived value and perceived trust were dropped whereas the perceived security, perceived ease of use and perceived

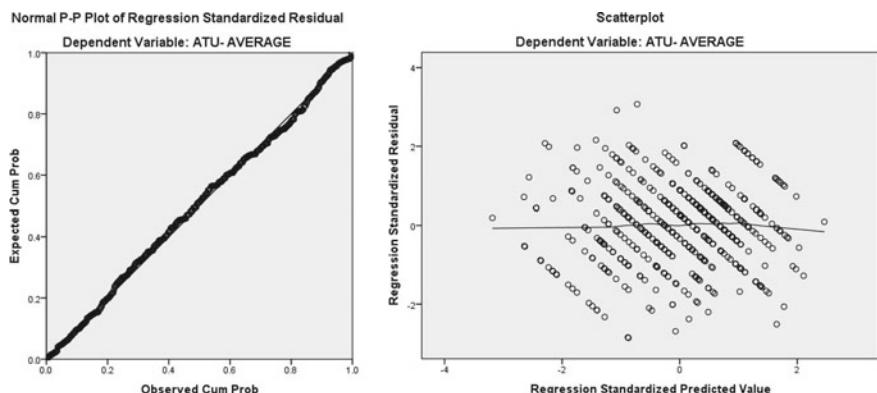
usefulness were found significant and thus retained ( $P < 0.001$ ). The final model had Durbin Watson value as 1.998 indicating no autocorrelation and the tolerance value are perceived security (0.543), perceived ease of use (0.539) and perceived usefulness (0.621) with the value of VIF being greater than 1 and less than 10 (1.610, 1.856 and 1.874) confirming no multi-collinearity. The scatter plot for dependent variable displays the relationship between predictors and the standardized residuals in which residuals are randomly scattered around zero and the losses curve displays almost linear relationship as shows in Fig. 1. The R<sup>2</sup> value thus measured for the model is 0.637 whereas the MSE value is 30.951. The perceived security has the highest standardized and unstandardized beta value (0.510 and 0.463). The best fit regression equation for the measurement of Attitude towards use of mobile wallets is as:

$$\text{ATU} = 0.255 + 0.093(\text{PU}) + 0.299(\text{PEOU}) + 0.510(\text{PS})$$

Here, Attitude towards use (ATU), Perceived usefulness (PU), Perceived Ease of Use (PEOU) and Perceived Security (PS).

Further for analyzing and comparing the measurement model with the Artificial Neural Network, the data was used to build Artificial Neural Network model and trained through multilayer preceptor algorithm. In the model, the Input layer has three significant independent variables which were adopted from the output of the multiple regression analysis. The output layer has one dependent variable and one hidden layer which consists of 6 hidden units for which the activation function used is sigmoid for all the layers [9]. The dependent variable attitude towards use is rescaled through normalized method. The error function used is the sum of squares.

To make the ANN results more robust and to avoid over-fit, ten-fold cross validation was implemented. Table 1 shows all the datasets used for training and testing in all 10 ANN models. The results are from the total neuron solution analysis using multilayer perceptron. The partition of the data set was commenced by assigning



**Fig. 1** Figure shows the residual PP plots and the scatterplot confirming normality of the data set

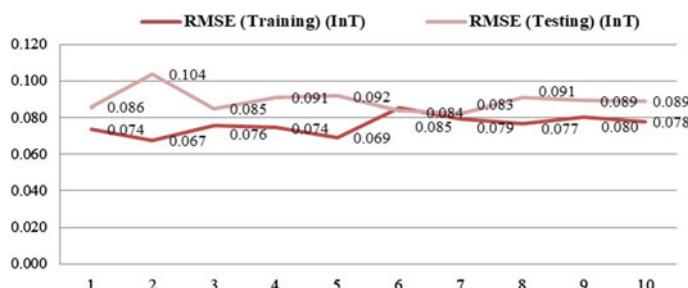
**Table 1** Sample dataset for testing and training

	Sample size (Training)(InT)	Sample size (Testing)(InT)	Holdout
N1	293	146	62
N2	299	145	56
N3	307	138	48
N4	307	165	55
N5	307	143	47
N6	308	134	58
N7	302	150	48
N8	311	128	59
N9	303	145	51
N10	311	126	55

60% for testing, 30% for training and 10% of the data was kept for holdout. Table 1 shows the data partition each time the ANN model has been run on the data set.

The computation of the sum of square error in testing and training datasets validates the ANN Model. The root mean square error (RMSE) value as close to zero points to a perfect fit model [10]. The average RMSE value obtained for the testing is 0.089 whereas the lowest is 0.083 and for training phase the average RMSE is 0.076 whereas the lowest is 0.067 which are within the reasonable limits thus making the model is accurate and reliable. Figure 2 shows the RMSE value which are within the limits for both the testing and training phases. The standard deviation for the training model is almost same as being 0.0052 and for the testing is 0.006. The R2 value (Table 2) for all the neural networks is between 0.70 and 0.75 and the average R2 value is 0.730 which is greater than the R2 value obtained in the Multiple Regression Analysis.

The sensitivity analysis in Table 3 shows the relative importance of the independent variables in the predictability of the dependent variables. The average normalized importance confirms that perceived security is the most important variable



**Fig. 2** Figure shows comparative RMSE value for the testing and training of the data set in ANN. The value of RMSE is below 0.1 which is within the desirable limits

**Table 2** Sum of squares, RMSE and R2

S. No.	Sum of square error (Training)	Sum of square error (Testing)	RMSE (Training)	RMSE (Testing)	R2
N1	1.588	1.082	0.074	0.086	0.754
N2	1.361	1.568	0.067	0.104	0.726
N3	1.76	0.994	0.076	0.085	0.738
N4	1.703	1.371	0.074	0.091	0.723
N5	1.47	1.206	0.069	0.092	0.747
N6	2.236	0.939	0.085	0.084	0.710
N7	1.897	1.021	0.079	0.083	0.738
N8	1.829	1.055	0.077	0.091	0.737
N9	1.96	1.155	0.080	0.089	0.708
N10	1.886	0.995	0.078	0.089	0.723
Average	<b>1.769</b>	<b>11.386</b>	<b>0.076</b>	<b>0.089</b>	0.730
SD	<b>0.2541</b>	<b>0.1968</b>	<b>0.0052</b>	<b>0.006</b>	–

The R2 value measures the variance in the dependent variable due to changes in the independent variables. Here, the higher value of R2 obtained through ten fold analysis of Neural Network being greater than Multiple Regression Analysis also indicate the greater significance of beta values

**Table 3** Normalized importance

Predictors	Importance	Normalized importance (%)
Perceived security	0.531	100
Perceived ease of use	0.332	62.42
Perceived usefulness	0.137	25.82

(0.531) in predicting the attitude towards use preceding perceived ease of use (0.332) and perceived usefulness (0.137). The beta values for the independent variables in the regression model also confirms to the relative importance.

### 3 Conclusion

The results of the study leads to the conclusion that the ANN works impressively in the classification of parameters selected for the study of mobile wallet adoption in comparison to the multiple regression analysis. The established ANN model has outperformed multiple linear regression model in predicting the consumer attitude towards use based on the R<sup>2</sup> value and the RMSE value. Further, in the predictability of the attitude towards mobile wallets use, the artificial neural network proved to be better in terms of examining the variance on dependent variable due to selected

independent variables. In terms of the relative importance of the independent variables in predicting the dependent variables, the ANN and regression provides similar results with PS, PEOU and PU being significant respectively. The model of ANN can be further redesigned by modifying and adding up the network parameters for dependent and independent variables. Additionally, ANN may be drawn on to study similar technology acceptance and adoption parameters. The accuracy of the predictive model can be enhanced with the variation in the sample units of the study. However this study is done for substantiation, the methodology proposed is applicable to any form of consumer adoption. Further the study can be conducted with structural equation modeling.

## References

1. F.D.A. Davis, *Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results* (Massachusetts Institute of Technology, Cambridge, MA, USA, 1985)
2. N. Koenig-Lewis, M. Marquet, A. Palmer, A.L. Zhao, Enjoyment and social influence: predicting mobile payment adoption. *Serv. Ind. J.* **35**, 537–554 (2015)
3. Y. Alsaawy, A. Alkhodre, M. Benaida, R.A. Khan, A comparative study of multiple regression analysis and back propagation neural network approaches for predicting financial strength of banks: an Indian perspective. *WSEAS Trans. Bus. Econ.* **17**, 627–637 (2020)
4. J. Lara-Rubio, A.F. Villarejo-Ramos, F. Liébana-Cabanillas, Explanatory and predictive model of the adoption of P2P payment systems. *Behav. Inf. Technol.* **40**(6), 528–541 (2021)
5. F. Liébana-Cabanillas, V. Marinković, Z.A. Kalinić, SEM-neural network approach for predicting antecedents of m-commerce acceptance. *Int. J. Inf. Manag.* **37**, 14–24 (2017)
6. G. Mohammadi, R. Tavakkoli-Moghaddam, M. Mohammadi, Hierarchical neural regression models for customer churn prediction. *J. Eng.* 9–15 (2013).
7. N.M.A. Bakar, I.M. Tahir, Applying multiple linear regression and neural network to predict bank performance. *Int. Bus. Res.* **2**(4), 176–183 (2009)
8. H.T. Pao, A comparison of neural network and multiple regression analysis in modeling capital structure. *Expert Syst. Appl.* **35**(3), 720–727 (2008)
9. Y.G. Kim, E. Woo, Consumer acceptance of a quick response (QR) code for the food traceability system: application of an extended technology acceptance model (TAM). *Food Res. Int.* **85**, 266–272 (2016)
10. K. Chen, A.H.S. Chan, Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). *Ergonomics* **57**, 635–652 (2014)
11. C.F. Chen, X. Xu, L. Arpan, Between the technology acceptance model and sustainable energy technology acceptance model: investigating smart meter acceptance in the United States. *Energy Res. Soc. Sci.* **25**, 93–104 (2017)
12. L.M. Badea, Predicting consumer behavior with artificial neural networks. *Procedia Econ. Financ.* **15**, 238–246 (2014)
13. A. Fadlalla, C.-H. Lin, An analysis of the applications of neural networks in finance. *Interfaces* **31**(4), 112–122 (2001)
14. E. Kalani, A. Elhami, R.B. Kazem-Zadeh, E. Kamrani, Selection of investment basis using neural networks in stock exchange. *Am. J. Ind. Bus. Manag.* **8**(03) (2018)

# Securing Unauthorized Access to Cloud Data Storage



Renu Devi, Ekta Narwal, and Sumeet Gill

**Abstract** In today's era, cloud data storage is a popular choice for data storage. Cloud technology gives computing and information storage services through the internet. Expansion in the measures of cloud users has raised protection and security concerns. Security of users' information was a major concern in cloud storage. This paper tried to simulate an experimental setup that provides a better security alternative for cloud architecture. We have applied the feed-forward backpropagation technique of artificial neural networks. The benefits of storing encrypted passwords are stored in the form of a weight matrix and which is almost impossible to crack.

**Keywords** Cloud computing · Artificial neural network · Feed-forward neural network · Data security · Back propagation algorithm

## 1 Introduction

With the advancement of technology, almost all users try to save their data in the cloud. With the help of the cloud, we can access our data from any device because the storage of data and its computing occurs on servers in the data center, rather than being locally stored in users' devices.

---

R. Devi · E. Narwal · S. Gill (✉)

Department of Mathematics, M.D. University, Rohtak, India

e-mail: [drsumeetgill@mdurohtak.ac.in](mailto:drsumeetgill@mdurohtak.ac.in)

R. Devi

e-mail: [renudevi.rs.maths@mdurohtak.ac.in](mailto:renudevi.rs.maths@mdurohtak.ac.in)

E. Narwal

e-mail: [ektanarwal.math@mdurohtak.ac.in](mailto:ektanarwal.math@mdurohtak.ac.in)

## ***1.1 Cloud Computing***

Cloud computing could be an innovation that increases storage capacity without investment in new infrastructure. Like every technology, the usages of cloud computing has its own merits and demerits. Data security from unauthorized access is a major concern we face nowadays [1].

Unauthorized access means when someone can access the data, application, system, and software without permission from the authorized user. Guessing passwords is an easy and common way to hack users' information by an unauthorized person. The passwords can be guessed manually or by using the software. They can access user names, passwords, and personal identification numbers (PINs) of authorized users.

## ***1.2 Artificial Neural Network***

Artificial Neural Network (ANN) includes the artificial neurons or nodes cluster, connected among each other. ANN methodology is dilated a few structures of the big Network of neurons in the brain of humans. The adjustable parameters of ANN are stated as weights that indicate the strength of a connection between 2 neurons or nodes. Once modification of the weights is done, ANN will acquire an arbitrary vector mapping from input to output. Majorly three layers are available: input, hidden, and output.

Many neurons are interconnected in the human brain. The artificial neural network works on the principle of a biological neural system. It is a useful tool for different applications which require broad characterization. The benefits of parallel processing in neural Network and their ability to data, depending on features, gives a promising stage for password storage in clouds. Artificial neural networks encompass a group of nodes and a set of connections, connecting pairs of nodes. Each connection contains a signal from one node to another, and an artificial neural network has a marked directed graph shape in which nodes carry out a few computations. Label considers the connection ability or weight specifying the level to which a connection increases or decreases signal. Distinct choices for the weights bring about different capabilities being evaluated by the Network. Weights of the Network are to start with random, and a learning algorithm is used to gain the values of the weights to accomplish the favored task. A neural network stores certain information to a problem in the weights of connections using a learning algorithm [2, 3].

### 1.3 Feed-Forward Neural Network

A feed-forward neural network constitutes processing components known as layers or subgroups. The network object is created in the first phase of training. It contains some layers which solve the complex relationship rapidly. Here, the signal travels in one way and thus, provides no feedback. Feed-forward ANN is that Network in which we use the input to obtain the output. The Network is being created by providing input and output values that initialize weight and bias values, and thus the size of the output layer is being determined. Through the training process, the weights are automatically adjusted so that the squared mean error between experimental and obtained results can be reduced. This network comprises three layers, i.e., the input layer, the hidden layer, and the output layer. The hidden neurons are computational elements of the hidden layer [4].

## 2 Literature Review

There have been a lot of research work applications using the artificial neural network. Researchers have used and still using these techniques for encryption and decryption processes. We have referred to many research articles to seek out the relevant data regarding the usage and applications of artificial neural network methodology. Here we present a brief of words related to our present work.

An increased security design using cryptographic techniques was developed by Singh [5] for cloud computing. They protected the data using ciphertext, retrieving the data and its key generation.

Multilayer neural networks and learning model proposed by Kumar et al. [6] to enhance privacy. The back-propagation algorithm of multiparty distribution in which the neural network can be trained without revealing the data to the users. They provide correctness and security analysis of these algorithms. The authors proposed this algorithm to preserve the privacy of the cloud server.

Negi et al. [7] have projected this model, which has adequate functionalities and capabilities to ensure data security and integrity using the encryption and decryption approach by using a back propagation neural network [7]. The projected model in the system provides help in building an extremely secure data security system for all the three layers of the cloud services offered to the cloud used by the cloud provider.

Mehtha and Saini [8] presented a model to review the security challenges, notably, the safety of data storage in the clouds by applying an erasable Correcting Code [8]. She has provided some offers to enhance the assurance of data storage within cloud computing systems.

A model for data security in the cloud, with the assistance of an Artificial Neural network, was projected by Suneetha et al. [9].

Elliot et al. [10] implemented a model for cloud security using Levenberg Marquardt Back Propagation (LMBP) algorithm. The training performance of ANN was verified and evaluated by the system's mean square error function [10, 11].

### 3 Experimental Setup

In the present simulation, we have taken the bit sequence of encrypted passwords that have come out to be 8 bits. Accordingly, we selected a neural network with 60 neurons in the hidden layer. The reason for choosing 60 neurons in the hidden layer is that if the number of neurons in the hidden layer is high, then the convergence of the training curve will be fast.

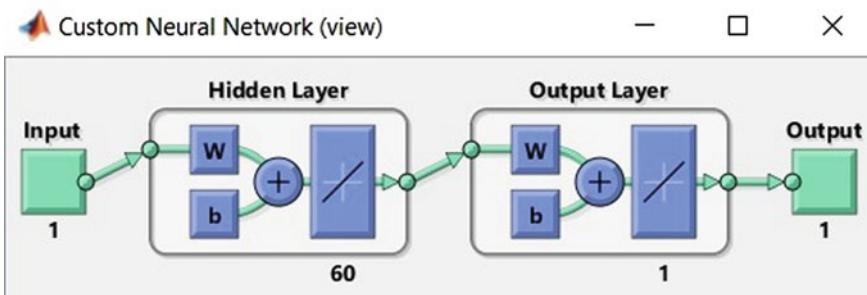
Figure 1 is the neural network view that is generated in MATLAB.

The sample data of encrypted passwords used for simulation is presented below (Fig. 2).

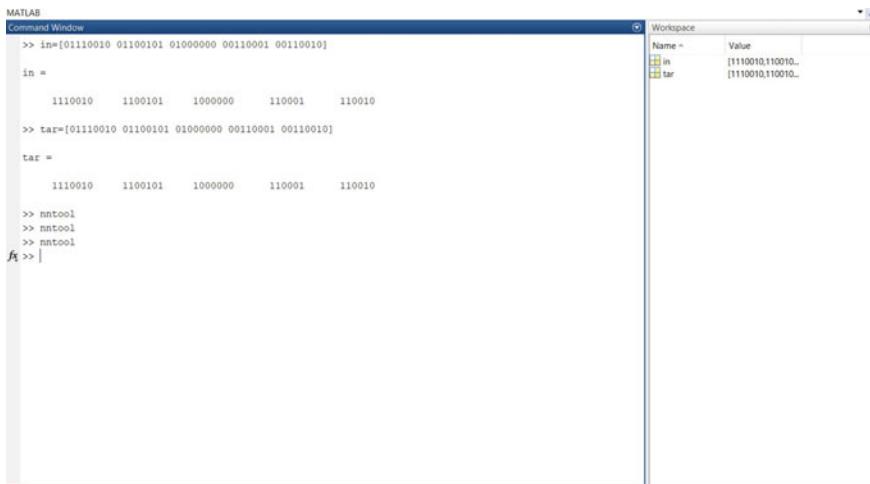
In this sample data, the input matrix is [01110010 01100101 01000000 00110001 00110010], and the target matrix is [01110010 01100101 01000000 00110001 00110010].

The algorithmic steps to conduct the simulations are as follows:

- Initially, the network is trained by providing data samples in the input layer. The target pattern is predicted with random test data in the input layer.
- The output value is given to the network, and if there is a difference between the experimental and desired result, then the mismatch value is redirected to the input layer so that weight and bias value are adjusted to reduce the error.
- At the next step, the neurons are passed to the hidden layer. The threshold function is used to get the output in the hidden layer. The function is given as  $1/(1 + \exp(-x))$  wherever  $x$  is a variable whose value is given by multiplying the weight vectors with the input vectors.



**Fig. 1** Neural network view



**Fig. 2** Screenshots of input and target data

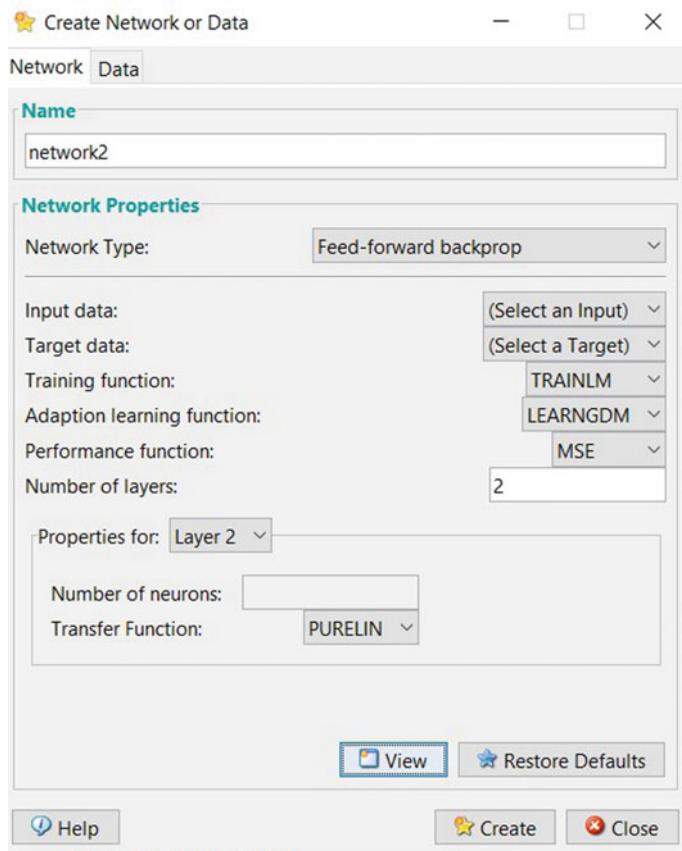
- The neurons at the hidden layer become input for the output layer and are again processed with threshold function. The output values are calculated from the output layer.
- At last, the experiment results were compared with the desired result. If there is any error, the same procedure is repeated until we get the range of tolerance factor.

## 4 Results and Analysis

We have done our simulations by using a feed-forward backpropagation algorithm to train the neural networks. In the experimental setup, we stored the network parameters and converted binary passwords into a weight matrix obtained from the networks after training. This weight matrix is almost impossible to crack.

We remodeled the above experimental setup to suit the training of our data sets. The following steps for the experimental setup are:

- Initially, we choose an encrypted password re@12 in the binary form [01110010 01100101 01000000 00110001 00110010] for the input layer in MATLAB, and the same is used the target for the Network.
- We use the feed-forward back propagation method, and after selection of input, output data, we choose the training function that is TRAINLM, adaption function is LEARNGDM, performance function MSE, 60 neurons, and PURELIN transfer function for better output performance (Figs. 3 and 4).
- After that, a regression of the neural network and the performance of our data is obtained.
- End.

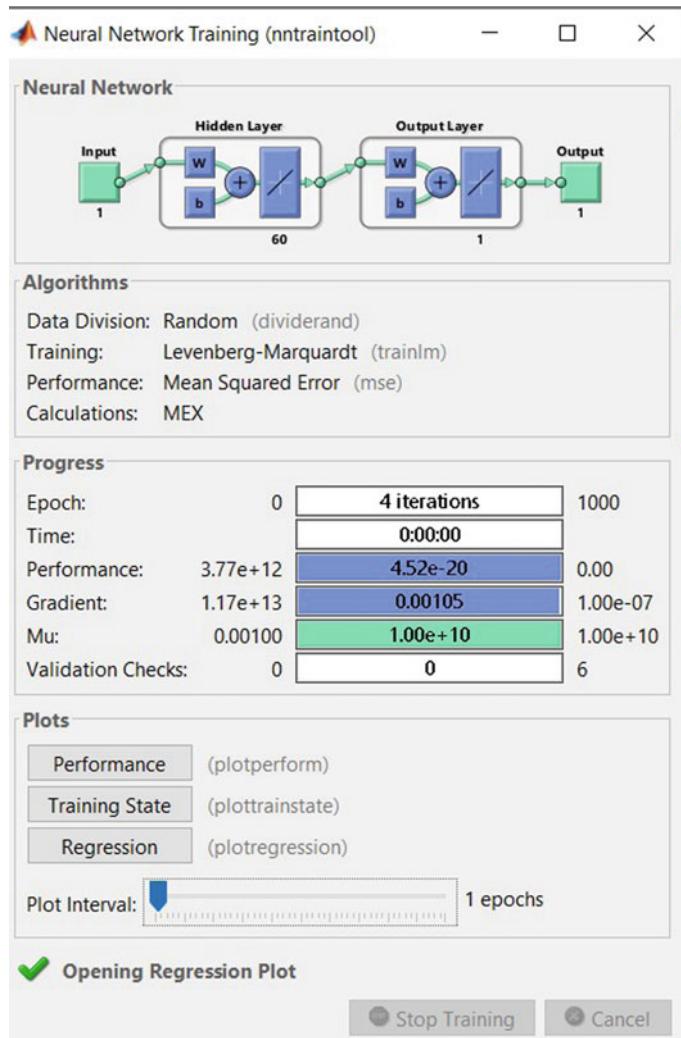


**Fig. 3** Screenshots of parameters to create network

The input and target data are loaded into the nn-tool with a  $1 \times 5$  dimension. Here, we have used 60 neurons in the hidden layer. The best validation performance is  $2.1684e-19$  at three epochs out of 4 epochs. The  $45^\circ$  line shows that network outputs (target) are equal to the input, which shows that our training is successful.

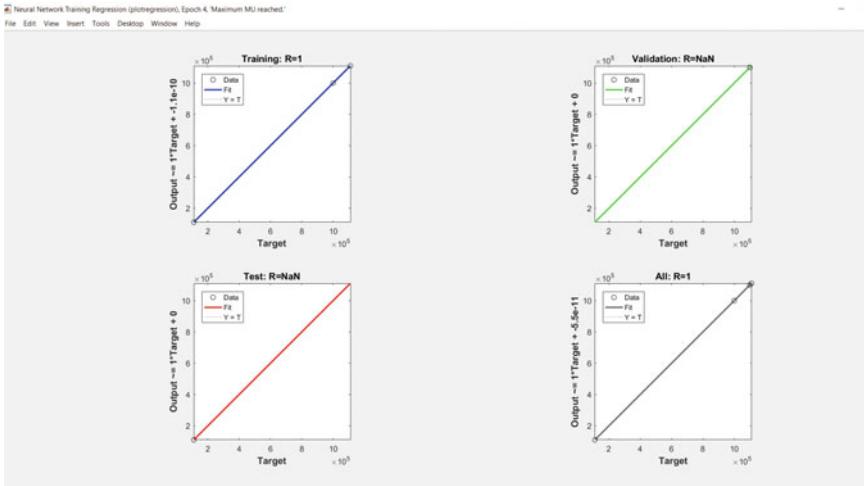
Figure 5 shows the regression of our sample data which shows a linear relationship between target and output. Here, the first plot represents the training of our data, the second is validation, and the third plot shows the test data. Here, the dashed line shows the perfect result that is  $\text{output} = \text{target}$ , and the R-value represents the relationship between output and desired result. Here  $R = 1$  indicates the exact relationship between target and output. If the value of R is very close to zero, then no linear relationship between output and desired results.

Figure 6 shows the convergence of data training, data validation, and data test of our sample data.

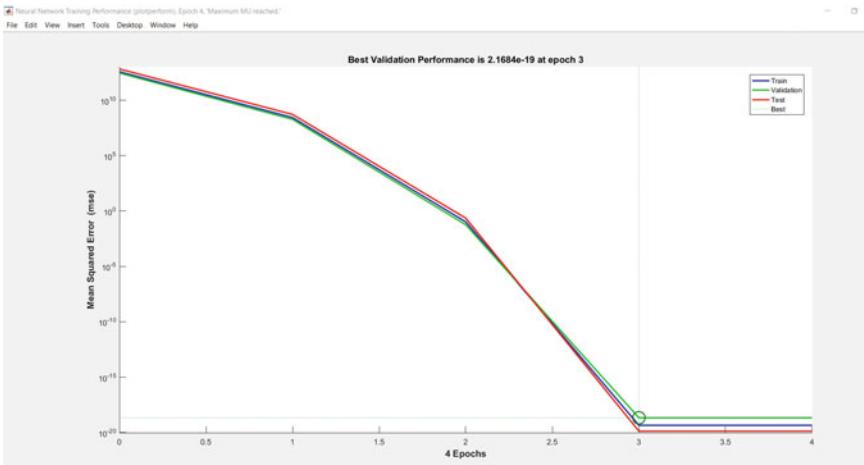


**Fig. 4** Screenshots of a trained network

When we train the network using the `trainlm` training function and `learngdm` adaption function, the network performance of our sample data converges at three epochs out of 4 epochs.



**Fig. 5** Regression plot



**Fig. 6** Performance of trained network

## 5 Conclusions

Cloud computing is one of the important technology for data storage. The major issue in cloud storage was unauthorized access to the user's information. Therefore, we had designed an experimental setup to provide better security based on the feed-forward back-propagation technique of ANN. The passwords are encrypted and trained into a network with an experimental setup in MATLAB. The results obtained through this experimental setup are analyzed and provide better security as the weight matrix

of the encrypted password is impossible to crack. Thus, this experimental setup will boost security for multiple applications.

## References

1. P. Srivastava, R. Khan, A review paper on cloud computing. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **8**(6), 17 (2018). <http://doi.org/10.23956/ijarsse.v8i6.711>
2. S. Joshi, A.N. Cheeran, MATLAB based back-propagation neural network for automatic speech recognition. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **3**(7), 10498–10504 (2014). <http://doi.org/10.15662/ijareeie.2014.0307016>
3. E. Narwal, S. Gill, Securing vehicle numbers using artificial neural networks. *Int. J. Comput. Sci. Eng.* **5**(10), 195–199 (2017). <http://doi.org/10.26438/ijcse/v5i10.195199>
4. P. Malik, S. Singh, B. Arora, An effective weather forecasting using neural network. *IJEERT* **2**(2), 209–212 (2014)
5. T.J. Singh, Double key encryption method (DKEM) algorithms using ANN for data storing and retrieval in cloud computing. *IOSR J. Comput. Eng.* **12**(2), 15–20 (2013). <https://doi.org/10.9790/0661-1221520>
6. P. Kumar, R. Muthu, D. Vijay, Privacy preserving back-propagation neural network in cloud computing. *IJARCET* **3**(6), 2086–2092 (2014)
7. A. Negi, M. Singh, S. Kumar, An efficient security framework design for cloud computing using artificial neural networks. *Int. J. Comput. Appl.* **129**(4), 17–21 (2015). <https://doi.org/10.5120/ijca2015906805>
8. S. Mehta, J. Saini, Tracking down unauthorized access by users in cloud, pp. 501–505
9. D. Suneetha, D. Rathna Kishore, G.G.S. Pradeep, Data security model using artificial neural networks and database fragmentation in cloud environment. *Int. J. Recent Technol. Eng.* **8**(2), 5972–5975 (2019). <http://doi.org/10.35940/ijrte.B3658.078219>
10. S.J. Elliot, V.I.E. Anireh, N.D. Nwiabu, A predictive model for cloud computing security in banking sector using Levenberg Marquardt back propagation with cuckoo search. *Int. J. Comput. Sci. Eng.* **7**(2), 42–47 (2020). <http://doi.org/10.14445/23488387/ijcse-v7i2p10>
11. S. Gill, E. Narwal, Modeling elliptical curve cryptography keys using back propagation algorithm. *Int. J. Innov. Technol. Exploring Eng.* **8**(9 Special Issue 3), 1525–1529 (2019). <http://doi.org/10.35940/ijitee.I3318.0789S319>

# A Survey on Water Usage Optimization Using Artificial Intelligence in Agricultural Sector



N. Bharathi and M. Hanumanthappa

**Abstract** Agriculture automation is a major source of concern and emerging subject across the world. Agriculture has a substantial contribution to the economy. The global population is quickly increasing, resulting in greater demand for food and labor. Traditional farming approaches were insufficient to achieve these goals. New automated procedures were created as a result. These new approaches met food demands while also creating employment opportunities for billions of people. The main purpose of this paper is to propose Artificial Intelligence (AI)-based water solutions in the agricultural sector. The goal is to highlight compelling new prospects for artificial intelligence (AI) interventions, to address the primary water management concerns that smallholder farmers confront. Water management may have a huge impact on their income and output, and it's ripe for technological innovation. By providing real-world examples, the topic of artificial intelligence and its applications in the water sector would be more accessible to a wider audience of readers. The examples shown here were chosen to exemplify how AI-based solutions may handle agricultural-related problems and deliver substantial advantages to the water industry.

**Keywords** Artificial intelligence · Irrigation · Water management

## 1 Introduction

### 1.1 Introduction About Agriculture

Agriculture consumes the most water globally; however, only a small portion of the water diverted for agriculture is actually utilized for agricultural development, with

---

N. Bharathi (✉)

KLE Society's S. Nijalingappa College, Rajajinagar, Bangalore, India  
e-mail: [ganesha.bharathi232@gmail.com](mailto:ganesha.bharathi232@gmail.com)

M. Hanumanthappa

Department of Computer Science, Bangalore University, Bangalore, India  
e-mail: [hanu6572@bub.ernet.in](mailto:hanu6572@bub.ernet.in)

the remainder being drained, evaporated, or runoff. Water shortage is becoming more prevalent around the world as a result of climate change and depletion, posing new issues for food production, processing, and preparation, as well as energy, industry, and other economic sectors, as well as our ecosystem and general well-being. Agriculture is both the cause and the victim of water scarcity, accounting for 70% of worldwide water withdrawal, making it a critical component in tackling the water problem.

Smallholder farmers, who own less than 2 ha of land and account for 80% of all farms worldwide, are among the world's most disadvantaged people, living in extreme poverty with a daily average income of \$3, and a poverty rate of more than 50%. Few technology breakthroughs aimed at addressing the issues encountered by smallholder farmers have had a significant influence on their livelihoods, with many losing steam in execution, delivery, and scale-up, or overlooking their most pressing needs as their communities and environs evolve. The market is flooded with technical solutions, but artificial intelligence (AI) has a huge untapped potential.

The main advantage of AI is its capacity to provide individualized services at scale and at lower rates, which is exactly what agriculture requires. Use of AI offers strong and feasible solutions that improve on many of the technology advancements, presently seen by analyzing the requirements of smallholder farmers, identifying the relevant levers in the system to address them, and measuring their potential routes to effect.

With this goal in mind, we identified water management as an area where we can make a significant difference in the lives of smallholder farmers. Water management can help smallholder farmers raise their income and yields, making it a major driver of enhanced profitability. Smallholder farmers, on the other hand, lack the knowledge and instruments necessary for effective water management at every stage of their trip, including water access, water consumption, and water replenishment. As a result, poor water management decisions are made, resulting in lesser income. The main aim of this paper is to bring to the forefront compelling new opportunities for AI interventions to address the major challenges smallholder farmers face in water management.

## ***1.2 Introduction About Artificial Intelligence***

Artificial intelligence (AI) is the simulation of human intelligence and will do the activities vision, speech recognition, decision-making, and language translation. Based on the data presented to them, AI systems seek to make intelligent suggestions or predictions. These inputs can take a variety of forms, including time-series data such as rainfall patterns, past yields from farmers, and photographs, among others. Similarly, the outputs can be recommendations or predictions in a variety of formats, such as rainfall forecasts for the following year, yield projections, or items recognized in an image.

## 2 Water Usage Optimization Techniques

We have prioritized two opportunities for AI in water management which ultimately helps farmers make better decisions that will improve their income and livelihoods:

1. Assess the water balance and improve crop planning using the Evapotranspiration method.
2. Optimize irrigation scheduling using smart Irrigation and dielectric methods.

### 2.1 Water Balance Using Evapotranspiration Method

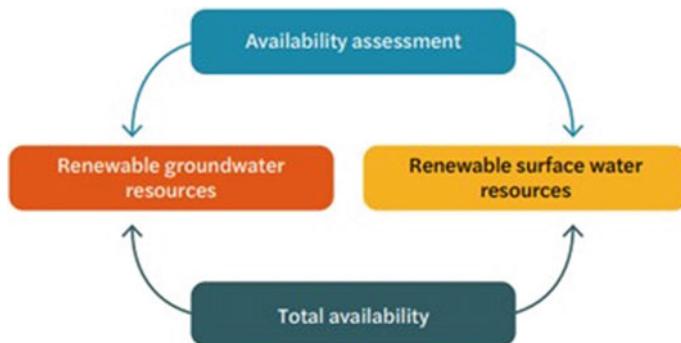
To aid the irrigation manager in making irrigation decisions, the state of the soil water for an irrigated crop must be monitored on a regular basis. Irrigation scheduling is usually done in one of two ways. To directly monitor soil water, one technique is to utilise soil moisture sensors. The soilwater balance technique, on the other hand, accounts for soil water in the rooting depth using weather data. Irrigation scheduling or water balancing strategies depending on weather or Evapotranspiration (ETc) are frequent terminology for this method.

#### 2.1.1 Creating Optimal Irrigation Scheduling and Distribution

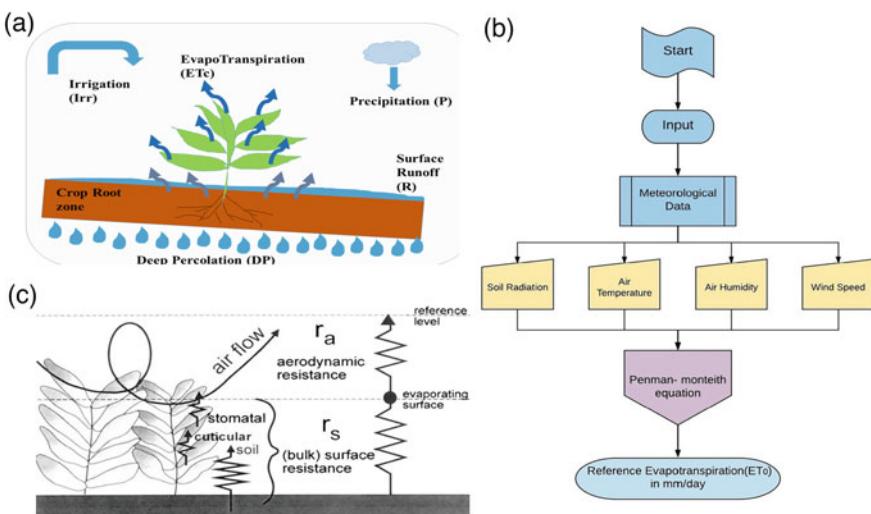
Farmers want to develop an irrigation strategy for their crops that maximizes output and quality while minimizing costs. Evapotranspiration has long been an important parameter for designing an irrigation system that is specific to a plant's demands. It is the product of evaporation from the land surface and transpiration from plants. Farmers can enhance their estimation of Evapotranspiration by using modern satellite images and weather predictions. Breakthroughs in the internet of things (IoT) sensor technology, on the other hand, can help inform far more precise irrigation decisions by detecting plant behavior instead of (or in addition to) soil and weather (Figs. 1 and 2).

### 2.2 Smart Irrigation

Smart irrigation technology is being developed to increase production without the use of a large number of humans by sensing water levels, soil temperature, nutrient content, and weather predictions. Hinnell et al. [1] proposed an irrigation system based on Neural Networks (Neuro-Drip). On the other hand, Goap et al. [2], Nawandar and Satpute [3] have presented an IoT-based smart irrigation management system. In collaboration with the Internet of Things, AI has made a substantial contribution in this field (IoT).



**Fig. 1** Process of calculating water balance



**Fig. 2** **a** Components of the soil water balance model for evapotranspiration, **b** flowchart for evapotranspiration, **c** FAO Penman-Monteith method

Smart irrigation necessitates gathering data on the amount of moisture in the soil, the water content of plants, the humidity in the atmosphere, temperature, and other elements. This data can be collected using soil moisture sensors, temperature sensors, humidity monitoring sensors, and other sensors.

A project called SWAMP written by Carlos Kamienski [4], which had explained the concept SWAMP, is used to develop IoT based methods and approaches for smart water management in irrigation system. By evaluating sensor data and other pertinent information, the SWAMP platform evaluates the water demands of individual farms, develops irrigation prescription maps, and provides them to the farmers, who then carry out the irrigation according to their preferences. The estimated water needs for

each farm are then sent into a component that optimizes water distribution management by delivering commands to actuators in devices that control water flow through canals until farmers may use them.

Shekhar et al. [5] In 2017, they created a robotic model that could detect the moisture content and temperature of Arduino and Raspberry Pi3 boards. The information is gathered on a regular basis and transferred to the Arduino microcontroller, which converts the analogue input to digital. The signal is transferred to the Raspberry Pi 3 (which includes the KNN algorithm), which then sends it to Arduino, which activates the irrigation system. Water will be provided based on demand, and sensor data will be updated and stored. Jha et al. [6] also used Arduino technology to create an automated irrigation system to save people and time in this field.

Another automatic watering approach was described by Varatharajalu and Ramprabu [7]. In this paper, they had discussed about soil moisture sensor, temperatures sensors etc., which helps in improving crop development.

One more method, subsurface drip irrigation it reduced water loss due to evaporation and runoff. Later, researchers built raindrop sensors that were operated via a wireless broadband network powered by solar panels in order to determine the demand for water supply to the fields. Using the GSM module, the raindrop sensor and soil moisture sensor transmit SMS to the farmer's phone, notifying them of the soil moisture content. As a result, the farmer can control the water flow by sending an SMS. As a result, we can expect this technology to identify sections of the field that require more water and prevent the farmer from watering when it rains.

### 3 Conclusion

We see tremendous potential for AI-based solutions to help tackle the underlying data difficulties and give more targeted, personalized, and relevant insights and recommendations to improve farmers' lives for each of the water management opportunities we've discussed. However, in order for any AI breakthrough to come to life and have a large-scale impact, we'll need an ecosystem that sets the ideal conditions for success, and we'll need to invest in and support a variety of initiatives and efforts to get there. A number of common enablers will make it easier for the AI approaches we've mentioned having the impact we want.

There are complementary areas of AI and irrigation exploration that lend further weight to this opportunity. Smartphone-based spectrometry for soil texture and nutrients, for example, is a new sector with a lot of potential. Fertigation, or the addition of fertilizer and mineral inputs to irrigation systems, is another developing subject in the irrigation industry, and one where AI has already proven its worth.

## References

1. A.C. Hinnell et al., Neuro-drip: estimation of subsurface wetting patterns for drip irrigation using neural networks. *Irrig. Sci.* **28**(6), 535–544 (2010)
2. A. Goap, D. Sharma, A.K. Shukla, C.R. Krishna, An IoT based smart irrigation management system using machine learning and open source technologies. *Comput. Electron. Agri.* **155**, 41–49 (2018)
3. N.K. Nawandar, V.R. Satpute, IoT based low cost and intelligent module for smart irrigation system. *Comput. Electron. Agri.* **162**, 979–990 (2019)
4. C. Kamienski, *Proceedings of the Global Internet of Things Summit* (IEEE, Piscataway, N.J., 2018), pp. 1–6. <https://doi.org/10.1109/GIOTS.2018.8534541>
5. Y. Shekhar, E. Dagur, S. Mishra, R.J. Tom, M. Veeramanikandan, S. Sankaranarayanan, Intelligent IoT based automated irrigation system. *Inter. J. Appl. Eng. Res.* **12**(18), 7306–7320 (2017)
6. K. Jha, A. Doshi, P. Patel, M. Shah, A comprehensive review on automation in agriculture using artificial intelligence. *Artif. Intell. Agri.* **2**, 1–12 (2019)
7. K. Varatharajalu, J. Ramprabu, Wireless irrigation system via phone call and SMS. *Int. J. Eng. Adv. Technol.* **8**(2s), 397–401 (2018)
8. M.J. Gebregiorgis, M.F. Savage, Soil-plant-atmosphere continuum research unit. School of Environmental Sciences, University of KwaZulu-Natal

# The Role of Artificial Neural Network in Word Sense Disambiguation (WSD)—A Survey



H. R. Roopa and S. Panneer Arockiaraj

**Abstract** The chore of reducing ambiguity in distinct sense of words is known as word sense disambiguation (WSD). It's a key area of research in computational linguistics to deal with the senses being assigned automatically to the words in a particular circumstances (Yuan et al. in Semi-supervised word sense disambiguation with neural models [1]). Human are naturally excellent at WSD and can tell the difference between senses utilized in the vocabulary through verbal language. On the contrary, computers have a hard time distinguishing between proper and incorrect meanings of words. Knowledge-based, Supervised, Semi-Supervised, and Unsupervised techniques have all been used to make progress in the problem of disambiguation. A better knowledge of human language will aid to computer performance in a variety of applications, including search and retrieval. The major goal of the paper is to describe a supervised neural network model that uses multiple strategies to maximize sense detection accuracy. The neural network's input layer will be made up of binary valued nodes based on whether or not frequently recurring context words connected to the ambiguous phrases are present. Amount of nodes in the outer layer will be equal to the amount of senses of the ambiguous word.

**Keywords** WSD · Neural network · Ambiguous word · Word embeddings

## 1 Introduction

Word Sense Disambiguation (WSD), which is a subset of Natural Language Processing (NLP) that deals with the issues of identifying the accurate meaning of a word in a given phrase. Various terms in the English language have multiple meanings or connotations. The challenge of determining the correct meaning is addressed by

---

H. R. Roopa (✉)

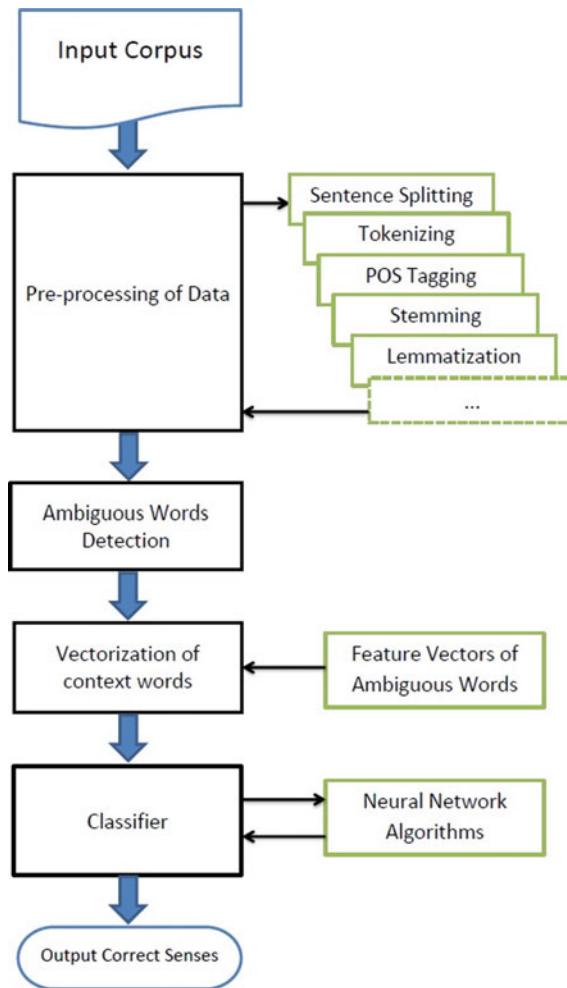
KLE Society's S Nijalingappa College, Rajajinagar, Bangalore 560010, India

e-mail: [roopahr0309@gmail.com](mailto:roopahr0309@gmail.com)

S. Panneer Arockiaraj

Computer Science Department, Periyar EVR College (Autonomous), Tiruchirappalli 620023, India

WSD. The answer to this challenge has an impact on search engine relevance. WSD is a skill, that the human mind excels at. A basic context is all that is required for people to comprehend the exact idea or meaning of a term. The logical ability of neural networks in the individual brain has resulted in the development of human languages. In computer science, developing the ability for computers to conduct processing the language with the same extent as human has been a long-term difficulty (Fig. 1).



**Fig. 1** Word sense disambiguation (WSD) using neural network

## 1.1 Artificial Neural Network (ANN)

A neural network is a collection of artificial neurons or nodes that form a network or circuit of neurons. ANNs are layers of the computing units which analyzes the input independently to replicate the functioning of the human brain. ANNs, like human brain learn via experience and exhibit gains in tasks as the amount of data available increases. ANNs with single or more concealed layers are known as Shallow Neural Networks (SNNs), whereas ANNs with multiple hidden layers are known as Deep Neural Networks (DNNs) (Fig. 2).

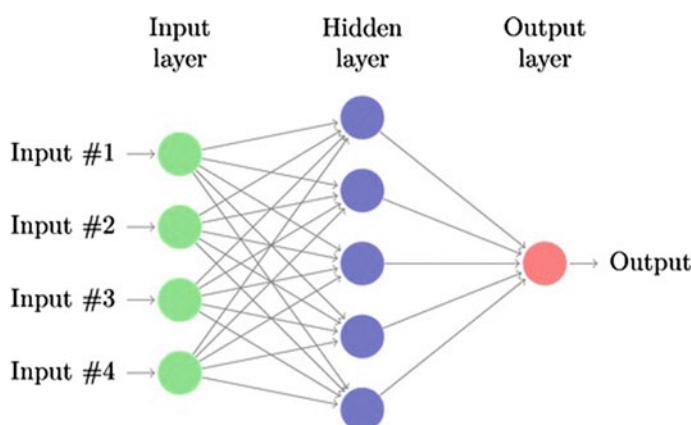
**Input Layer:** This layer receives inputs from the external world to the model to learn and draw conclusions. It is also referred as input nodes. The inputs from these nodes are passed to the next layer called hidden layer.

**Hidden Layer:** Inputs from the input layer are passed to set of neurons for computation. The number of hidden layers in a neural network is not restricted. A simple network consists of a single hidden layer.

**Output layer:** Once the computation is performed, conclusion is drawn from the output layer. Nodes in the output layer can be one or more in number based on the classification problem. For a binary classification problem the output node is 1 where as in case of multi classification problem, output nodes can be more than 1.

## 2 Different Approaches of WSD

WSD (Word Sense Disambiguation) approaches have been investigated and researched extensively in the past. The different approaches of WSD are:



**Fig. 2** Simple neural network

- Knowledge Based Approach
- Supervised Approach
- Semi-Supervised Approach
- Unsupervised Approach.

## ***2.1 Knowledge Based Approach***

To determine the correct sense of words, algorithms based on knowledge utilize lexical resource such as Machine Readable Dictionaries (MRDs) and WordNet. These Algorithms are simple to use and were the first to be created while attempting to address the WSD challenge. To begin the process of disambiguation, a knowledge based system just needs access to dictionary resources. The performance of the algorithm is restricted by the speed with which they can search for and retrieve these resources. Latency increases as the size of the resources grows, and performance suffers as a result [2].

## ***2.2 Supervised Approach***

The term “supervised technique” comes from the fact that it necessitates human intervention. To make supervised models to function as predicted, a large amount of labeled data is required. The greater the data set supplied, the better the system’s forecast accuracy. A defined learning set is created to allow the system to guess, true meaning of an ambiguous words based on a few texts that each has a distinct meaning for that word [2]. Based on the defined learning set, system determines the true sense of an ambiguous word in a certain situation. Supervised strategy consistently outperforms all other methods.

## ***2.3 Semi-supervised Approach***

Many WSD algorithms employ semi supervised approach, which is a hybrid of supervised and unsupervised learning methods. It accepts both labeled and unlabeled data, making them beneficial when training data is scarce. For each word, the boot strapping processes begin with a little quantity of data. A small quantity of tagged or labeled data is utilized to guide the primary classifier using any of the supervised approaches. After then, unlabeled data is supplied to the classifier in order to extract a larger labelled dataset with only flawless classification. Typically, such methods are iterative, with each iteration involving training on a larger dataset. The resulting data set grows in size till the end of the process after a particular amount of iterations or until the dataset’s upper limit size is achieved.

## 2.4 *Unsupervised Approach*

For WSD researchers, unsupervised learning approaches are the most difficult to deploy. We, mainly mean that meanings of the words can be derived using additional similar phrases when we use unsupervised methodologies. The idea behind unsupervised learning is that words with similar meaning are surrounded by similar words. The goal is to classify new occurrence into derived clusters, and word senses are created by constructing clusters of word occurrences. This approach finds clusters rather than assigning sense labels.

## 3 Supervised Approach to Resolve WSD

Various approaches to resolving word sense ambiguity have been proposed. The supervised method to WSD has been shown to yield the most accuracy among all the approaches. As a result, a supervised technique is applied in the proposed model. Artificial Neural Networks replicate the environment in which an ambiguous word occurs and the neural network should be able to correctly forecast the ambiguous word's right meaning. Because this type of model is supervised, it necessitates a large amount of labeled data in order to create an accurate neural net classifier.

### 3.1 *Word Embeddings*

The neural network's input feature vectors for an ambiguous word will be produced by making use of its word embeddings. Neural word embeddings are the vectors that are utilized to represent words. Word embeddings are made up of words that are similar and regularly used in context. They can be made using a variety of methods, including Word2Vec, recurrent RNN models like LSTM, and so on. Word Embeddings calculates a binary value for the input, i.e. no match is represented by 0 and total match is represented by 1.

### 3.2 *Feature Vectors*

Dataset input and word embeddings will be used to build input vectors for the neural network. If embedded terms are present in the input framework, they will have a value of '1' in the feature vector, otherwise a value of '0'. After each word in the input context is confirmed for its existence in word embeddings, the produced feature vector will be passed to the neural network input layer. As a result, the input

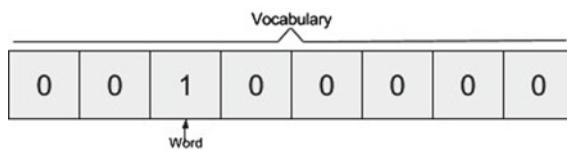
layer of the Neural Net will have the same number of nodes as the number of word embeddings.

Consider an example, let the input sentence be “THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG”.

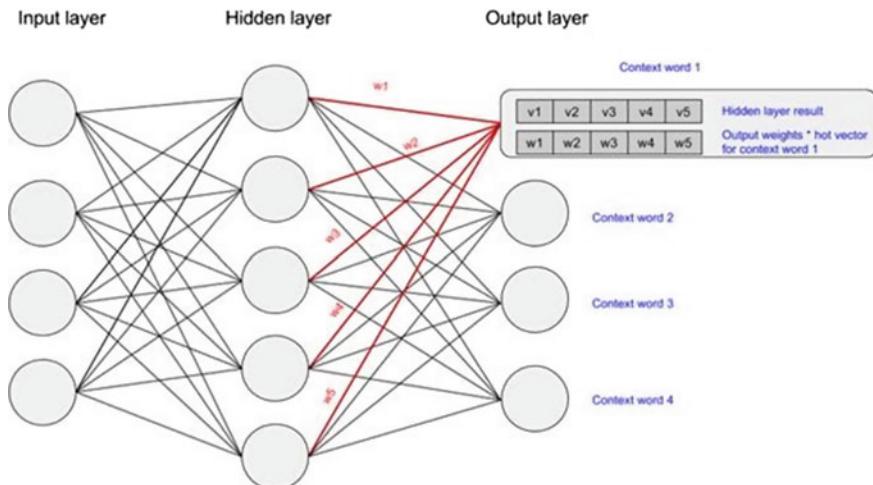
The vocabulary words in the given input sentence are THE, QUICK, BROWN, FOX, JUMPS, OVER, THE, LAZY, DOG. The word ‘BROWN’ is represented in the vector form as: [0 0 1 0 0 0 0 0 0].

The input vector given to a neural network with a single hidden layer is as shown in Fig. 3.

The number of nodes in the outer layer of the neural network is equal to the number of alternative senses for the ambiguous word. The projected sense will be represented by the node with the maximum integer value calculated among the several output nodes [3]. If maximum value is found in the third node of the output layer, the system has anticipated the third sense (Fig. 4).



**Fig. 3** Input vector with a single hidden layer



**Fig. 4** The state of the context word's output layer

## 4 Conclusion

In this paper, various approaches to resolving word sense disambiguity have been proposed. The four techniques, knowledge based approach, supervised approach, semi supervised approach, and unsupervised approaches are used by majority of the systems. The supervised method to WSD has been shown to yield the most accurate of these approaches. As a result, a supervised approach is applied to the proposed model. Artificial Neural Network replicate the environment in which ambiguous words occur and neural network should be able to correctly forecast the ambiguous word's right meaning. Because this type of model is supervised, it necessitates a large amount of labeled data in order to create an accurate neural net classifier. Also, this paper suggested a WSD model based on neural network with a goal of maximizing accuracy for a specific natural language processing job.

## References

1. D. Yuan, J. Richardson, R. Doherty, C. Evans, E. Altendorf, Semi-supervised word sense disambiguation with neural models (2016)
2. C.D. Kokane, S.D. Babar, Supervised word sense disambiguation with recurrent neural network model. Int. J. Eng. Adv. Technol. (IJEAT) **9**(2) (2019)
3. M. Lesk, *Automatic Sense Disambiguation Using Machine Readable Dictionaries: How to Tell a Pine Cone from an Ice Cream Cone*. Bell Communications Research, Morristown, NJ (1986)

# Covid-19 Challenges for Digital Education and Critical Analysis



Puninder Kaur, Mehak Giri, Rajwinder Kaur, Avinash Sharma, and Manik Bhan

**Abstract** Online study is not a new technology it was used since 2013 but become popular during the Corona virus. This research articles contain the study of the comparison between study through virtual platform and study by physical medium. Our study involves how the both traditions are different from each other. In this review paper we will also discuss the different online tools and methods used by the particular organisation for taking online classes. The advantage and disadvantage of online education and how it has changed the perspective of learning are also discussed. We will also discuss the importance of campus learning for improving the social skills and technical skills.

**Keywords** Remote learning · Challenge · Virtual platform · Face-to-face learning · Covid-19

## 1 Introduction

Online learning is the education taken through different virtual platform like zoom, MS teams, Google meet and so on. E-learning suddenly became popular due to the deadly and infectious disease known as covid-19. Corona virus has deeply affected the education system. Earlier students used to take class through physical medium but now the situations are totally changed. Pandemic has forced educator to adapt online teaching learning process [1]. Online learning has been introduced as the substitution of face-to-face learning [2]. Maintaining the social distancing is mandatory in all schools, colleges, universities even at every place. Therefore, they discontinued inperson teaching and have to adapt online learning. Adapting this new tradition was challenging for both students as well as teachers. As Covid was sudden so, most of the

---

P. Kaur (✉) · M. Giri · R. Kaur · M. Bhan

Chitkara Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India  
e-mail: [puninder.k@chitkara.edu.in](mailto:puninder.k@chitkara.edu.in)

A. Sharma  
Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, India

faculty members facing the challenges such as lack teaching skills, early preparation such as make PPT for their session, and lack of knowledge of educational software [3]. During covid fear, anxiety and negative thoughts was so high it was difficult for everyone to focus on their targets as a result the performance of students become low [4]. Cyberchondria was one of the challenged faced by people [5]. Platforms like zoom, teams and google meet has grown exponentially during these days and other video meeting tools that gives the better improvement of screen sharing as a whole or a particular page that makes the best virtual meeting, online sessions, webinars, and so on. Most of the organizations readily recommended it to their students, staff, and other faculty with efficient price [6]. There are various online meeting platforms are available some are paid and some are free of cost. The List of particular platforms are shown below (Fig. 1):

Online learning platform allowed the experts to conduct and attend virtual sessions anybody around the world. Therefore, anyone can access and gain their knowledge from anywhere [7]. Despite of this facility provides by online platform it carries a stigma a stigma of being lower than physical learning [8]. In physical learning body language, eye contact, and physical gestures and teaching aids are significant tools but in online only their voice and vocal functions are the tools only [9]. Teaching should engage the students with learning activities which includes fun for student's despite of the platform [10]. During online classes the most common problems occurred were



**Fig. 1** Online meeting platform

their internet connection speed, network quality and also the desktop problem. One economical problem that was faced by people was high bill of internet and many health issues such as negative effects on eyes, back ache and so on [11]. During corona virus practical classes, laboratory work, library visit, peer tutoring, remedial teaching, research were put on hold [12] that's why Virtual learning was not capable to replace Lesson which contains logic building more practical experiment [13]. Online education is still out of the reach of many students, who have limited or no access to computers and internet connectivity at home. Especially at the border areas internet facility are very low, hence such students face many problems during online classes and e-exams. During pandemic many have lost job and can't afford laptop and internet connection for their children to attend online classes.

There are numerous advantages of e-learning line anyone can be continue through a training program at their own place at flexible timing and also it is cheaper as any of your content are immediately available to your learning audience across the world [14]. During online classes anxiety disorder, social anxiety disorder, panic disorder and obsessive-compulsive disorder increased among students due to the less physical activity involved to relieve the stress of students [15]. But now, the fact is that in education quality terms face-to-face interactions between teachers and students can't be replicable by online learning. As campus learning improves once social skills speaking ability, encourages students to participate in healthy debates. While taking or joining classes educators and students suffer following challenges (Fig. 2):

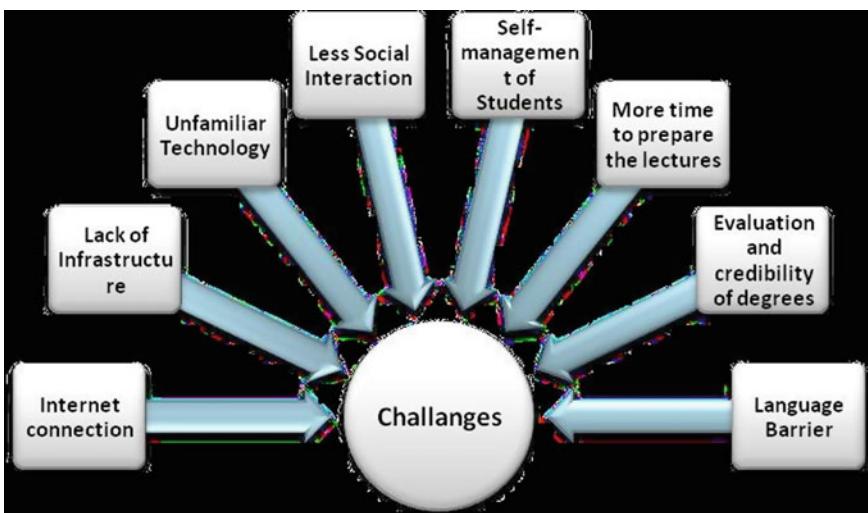


Fig. 2 Challenges of online study [18]

## 2 Review of Literature

This online study is helps the students as well as teachers to continue their study during the pandemic situation. The e-study firstly starts in china in March, 2020. Most of the online study is started in developing countries during lockdown because the developed countries are already working on this platform. Most of the distance courses are already run their study via online platform such as google meet, zoom, webex, MS Team and so on. Various authors contribute their efforts to provide us information about tools, methods, advantages and future scopes their articles. The brief of the literature are mentioned below (Table 1):

## 3 Conclusion

In conclusion, I would like to say both the medium are good if the students try to learn new things. We cannot neglect the fact such as network error, lack of telecommunication devices and system failure etc. These are the main barrios in online learning methodology. Even some students don't attend the virtual classes attentively as there is less strictness. On other hand, others students who wants to gain knowledge and learn something new. Therefore, they are active in both class whether it is online or offline. In future, online learning will for sure be the one of the most comfortable means of learning if the network quality is improved along with reasonable network package.

**Table 1** Literature review

Author	Focus	Tools/method	Advantage	Issues
Lockee [1]	Online education in the post-COVID era	Hyflex model	Remote teaching	Accessibility, challenges and opportunities using online platform
Bao [2]	Comparison online learning and offline learning	E-platform	Comfortable	The implementation of remote classes is high
Fatonia et al. [3]	This research articles focus on impact on shifting study mode from online to offline	Major five principles are provided to improve the online teaching methodology	The best part of this article is to improve the standard of online educations	Only emphasis on five principles but types of platform is still pending
Sands et al. [7]	Impact of e-learning	Outlook	Online and distance learning become flexibility	For distance learning a great application
Serhan [8]	Adaption of Zoom	Zoom	Easier interaction	For workshops its useful
Mahmood [9]	This research articles focus on different instructional strategies for improves the overall results of the students	Use different tools and strategies to overcome the problems faced by teachers as well as students	Provide training to the persons involved in online teachings	This remote learning facility is just implemented for the developing countries only
Beatty [10]	Ideal learning purpose	Hybrid-flexible course design	Accessible to anyone	New platforms for physical learning
Putri et al. [11]	Time paper focus on the impact on teachers, parent and student after joining online classes	Different online tools	This paper firstly provides training of particular tools to adopt	This paper doing research only in two areas
Mishra et al. [12]	Online learning	Ted-Ed, Coursera, Google Classroom, Bakpax, and so on	Gives the information of using all methodologies	The data set can be increased
Hodges et al. [13]	Comparison between emergency remote teaching and online learning	Virtual conferencing	Easy to access	Will find new audience

(continued)

**Table 1** (continued)

Author	Focus	Tools/method	Advantage	Issues
George [14]	This paper is mainly discuss about the challenges and opportunities while using remote learning method	Google Classroom	Educational innovation	Gateway to new opportunities for educators
James [15]	Advantage of elearning	All platforms	Anyone can proceed	For remote user it will always remain beneficial
Chogyel et al. [16]	In these articles various online platforms and electronics devices are discussed	Parallel mixed method design	Various platforms are used but google classroom are most effective for secondary schools	This study focus on the problems faced by the students only but problems faced by the teachers are still pending
Mertens et al. [17]	Online study	All social platform		After some time fear of covid will be overcome

## References

1. B.B. Lockee, Online education in the post-COVID era. *Nat. Electron.* **4**(1), 5–6 (2021)
2. W. Bao, COVID-19 and online teaching in higher education: a case study of Peking University. *Hum. Behav. Emerg. Technol.* **2**(2), 113–115 (2020)
3. N.A. Fatonia, E. Nurkhayatic, E. Nurdiawatid, G.P. Fidziahe, S. Adhag, A.P. Irawanh et al., University students online learning system during Covid-19 pandemic: advantages, constraints and solutions. *Syst. Rev. Pharm.* **11**(7), 570–576 (2020)
4. A. Farooq, S. Laato, A.N. Islam, Impact of online information on self-isolation intention during the COVID-19 pandemic: cross-sectional study. *J. Med. Internet Res.* **22**(5), e19128 (2020)
5. M.J. Paschal, D.G. Mkulu, Online classes during COVID-19 pandemic in higher learning institutions in Africa. *Global Res. High. Educ.* **3**(3), 1–21 (2020)
6. A. Hampshire, W. Treder, S. Chamberlain, A. Jolly, J.E. Grant, F. Patrick et al., Cognitive deficits in people who have recovered from COVID-19 relative to controls: an N = 84,285 online study. *EClinicalMedicine* <https://doi.org/10.1016/j.eclim.2021.101044> (2020)
7. T. Sands, F. Shushok, The COVID-19 higher education shove. *Educause Rev.* (2020) (available at <https://er.educause.edu/blogs/2020/10/the-covid-19-higher-education-shove>)
8. D. Serhan, Transitioning from face-to-face to remote learning: students' attitudes and perceptions of using Zoom during COVID-19 pandemic. *International Journal of Technology in Education and Science (IJTES)* **4**(4), 35–342 (2020)
9. S. Mahmood, Instructional strategies for online teaching in COVID-19 pandemic. *Hum. Behav. Emerg. Technol.* **3**(1), 199–203 (2021)
10. B.J. Beatty (ed.) *Hybrid-Flexible Course Design*, Ch. 1.4 (2019) (available at: <https://edtechbooks.org/hyflex>)

11. R.S. Putri, A. Purwanto, R. Pramono, M. Asbari, L.M. Wijayanti, C.C. Hyun, Impact of the COVID-19 pandemic on online home learning: an explorative study of primary schools in Indonesia. *Int. J. Adv. Sci. Technol.* **29**(5), 4809–4818 (2020)
12. L. Mishra, T. Gupta, A. Shree, Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *Int. J. Educ. Res. Open* **1**, 100012 (2020)
13. C. Hodges, S. Moore, B. Lockee, T. Trust, M.A. Bond, The difference between emergency remote teaching and online learning. *Educcause*
14. R. George, Impact and consequences of covid-19 in educational sector: challenges and opportunities (2021)
15. G. James, Advantages and disadvantages of online learning. Retrieved July 1, 2006 (2002)
16. N. Chogyel, N. Wangdi, Y. Dema, Evaluating the challenges in online learning during the COVID-19 pandemic in a middle secondary school. *Int. J. Didactical Stud.* **2**(2), 101459 (2021)
17. G. Mertens, L. Gerritsen, S. Duijndam, E. Salemink, I.M. Engelhard, Fear of the coronavirus (COVID-19): predictors in an online study conducted in March 2020. *J. Anxiety Disord.* **74**, 102258 (2020)
18. G. Rani, P. Kaur, T. Sharma, A. Sharma, Digital education challenges and opportunities, in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)* (IEEE, 2021), pp. 197–202

# Cloud Computing: Concepts and Techniques with Service Provider Analysis



Himanshu Barhaiya and Nilesh Jain

**Abstract** Cloud computing conveys IT-related capacities as associate degree administration through the net to various shoppers and these administrations are charged enthusiastic about utilization. Several distributed computing suppliers, for instance, Google, Microsoft, Yahoo, IBM, and Amazon are moving towards the reception of cloud innovation prompting an intensive intensifying within the use of various cloud administrations. Amazon is that the pioneer during this field as a result of its progressive variety of style highlights contrasted with others. To handle the problems of cloud specialist co-ops and shoppers totally different open supply devices and business apparatuses are being created. Despite the very fact that heaps a lot of advancements have occurred in distributed computing zone, various difficulties, for instance, security, ability, plus coming up with, virtualization, so forth are however to be adjusted. This paper surveys distributed computing the worldview as way as its verifiable advancement, ideas, innovation, apparatuses, and totally different difficulties.

**Keywords** Cloud Computing · Service Provider Analaysis

## 1 Introduction

The term ‘cloud’ was begat from the laptop system charts that use it to hide up the unpredictability of foundation enclosed. Distributed computing is finding out a fantastic breadth towards IT ventures [1], scholastics, and individual shoppers seeable of its usability, on-request access to rearrange assets, negligible administration toil, and diminished expense. This new Cloud computing innovation has typically unfolded within the market Associate in Nursing there’s an increment within the range of undertakings [2]. It is charming the cloud shoppers by giving advantages at an occasional price, pay for-use technique, disseminated nature, fast conveyance

---

H. Barhaiya · N. Jain (✉)

Department of Computer Application, Mandsaur University, Mandsaur, India  
e-mail: [hodca@meu.edu.in](mailto:hodca@meu.edu.in)

of process assets and furnishes info reposition focus with infinite area and unimaginable problem-solving limit with relevance swing away and overseeing info. The cloud model contains of 5 basic attributes—wide system gets to, quick skillfulness, plus pooling, on-request self-administration, Associate in Nursing calculable administration; 3 administration models—programming as an administration (SaaS), stage as Associate in Nursing administration (PaaS), Associate in nursing framework as an administration (IaaS); and 4 arrangement models—open cloud, non-public cloud, mixture cloud, and network cloud.

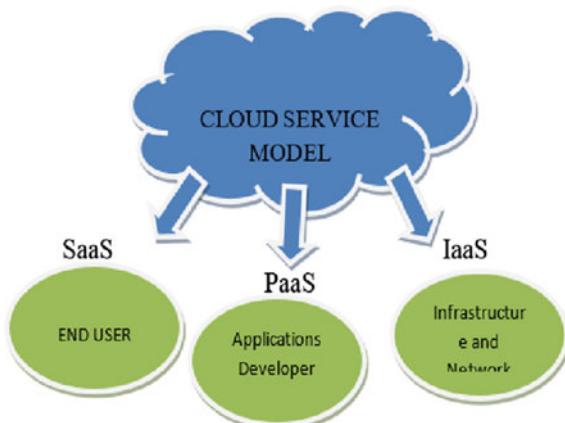
## 2 Cloud Service Models

### 2.1 Software as a Service (SaaS)

The method for conveyance of title application as associate administration on the net is observed as package as a service. Rather than introducing the merchandise on his computer, the consumer will simply access it by suggests that of the net [3]. It makes the consumer free from overseeing complicated programming and instrumentation. SaaS shoppers do not have to get programming or instrumentation, keep up, and update. the most issue the consumer should have an internet association and at the moment entrance to the appliance is straightforward.

Model, Microsoft workplace 365, Google Apps, and so on (Fig. 1).

**Fig. 1** Cloud services model



## 2.2 *Platform as a Service (PaaS)*

An advancement domain or platform is given to the consumers as a service in PaaS, whereat a consumer will send their own product and cryptography. The consumer has the liberty to develop his own applications which will keep running on the supplier's foundation [3]. Item as specialist organizations offers a predefined creation of a operating framework and application server to induce the administration limit of the applications. as an example, LAMP (Linux, Apache, MySQL, and PHP), J2EE, Ruby, and so on.

## 2.3 *Infrastructure as a Service (IaaS)*

Many computing resources area unit given by the IaaS as capability, arrange, operating framework, equipment, and capability gadgets on interest. IaaS purchasers will get to the administrations utilizing a good zone prepare, for instance, the online [3]. For example, a consumer will create virtual machines by login to the IaaS stage.

## 3 Comparison About Cloud Services Model (CSM)

See Table 1.

## 4 Analysis of CSM (Cloud Services Model)

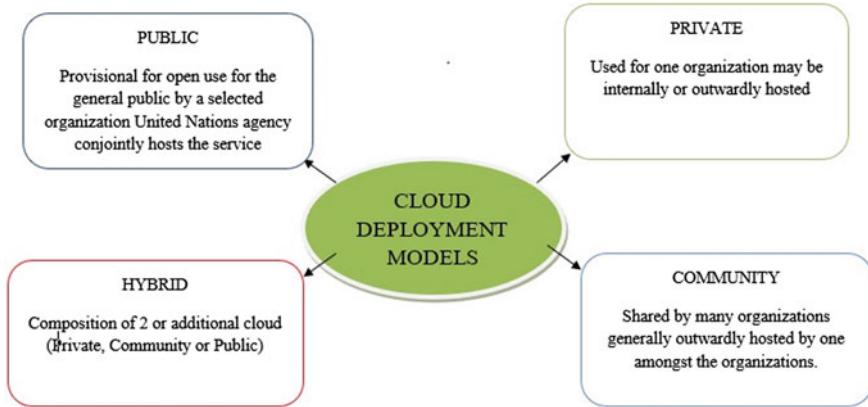
IT businesses face problems such as rising costs and increasing Constant need for capital investment; sprawl of servers; growing need for capital investment Operational costs, energy costs and requirements are ballooning, Increased system complexity and the need for advanced training Talent to help them, use of variable redress, demand for Simplification of company processes, time to consumer pressures.The adoption of IT or cloud computing as a utility service.

## 5 Cloud Deployment Model

As cloud innovation is giving shoppers such Brobdingnagian numbers of benefits, these blessings should be organized keen about the client's necessity. The Cloud readying model speaks to the correct category of cloud conditions keen about possession, size, and access and what is more portrays the character and motivation behind

**Table 1** Comparison about Cloud Services Model (CSM)

Attributes	PaaS (Platform as a Services)	SaaS (Software as a Services)	IaaS (Infrastructure as a Services)
Service providers	Azure, Google app. engine, Net suite	Salesforce.com, Office lives, Google apps	IBM, Microsoft, Google, Amazon
Application management	By the application developer	By the end user	By the vendor
Data management	By the application developer	End user	Vendor
Runtime management	Supplier	End user	Vendor
Operating environment	Total operating environment is included e.g. Windows, .Net, Linux, J2EE	Operating environment is largely irrelevant and fully functional apps provided e.g CRM (Customer Relation Management), E-mail, ERP (Enterprise Risk Management)	Virtual platform on which required operating environment and applications can be deployed and storage space
Number of providers	Few cloud platforms	Thousands of applications in the cloud	Elite group providers
Types of services	Integration as a Service	Dynamic services	Infrastructure
Movement data	Source code	Content and business processes	Operating system or virtual machine
Client control	The consumer has management over the deployed applications and configuration settings for the app. hosting environment	The shopper has restricted management of the user-specific app. configuration settings	The shopper has management over o/s, storage, and deployed applications
Focus	Developers, development tools and environments, social media	Services consumer, applications including archive, backup, ecommerce, social media	Virtual resources, servers, storage, networking, hardware and software services
Access type	One will access tools to assist write associated deploy an application, supported technology owned and managed by somebody else	One will use, access applications and storage facilities while not having to transfer or manage package on your computer	One can use the storage, computer resources and networks owned and managed by someone else



**Fig. 2** Cloud deployment model

the cloud. Most associations actualize cloud foundations to limit capital consumption and management operating expenses. To know that causing model matches your requirement and need it's essential for shoppers even as students to understand the four sub-classes of models for readying.

### 5.1 *Public Cloud Model*

The Public cloud may be a registering administration provided by the outsider suppliers on the open net [4]. These administrations area unit accessible for any shopper World Health Organization must utilize them and that they ought to pay only for the administrations they gone (Fig. 2).

### 5.2 *Private Cloud Model*

The computing services given over the online or non-public system go beneath the non-public cloud and these administrations area unit offered clearly to the chosen purchasers rather than everyday voters [4, 5]. Higher security and protection are selected by non-public mists through the firewall and interior facilitating.

### **5.3 *Hybrid Cloud Model***

Hybrid cloud may be a mixture of public cloud and personal cloud. within the half and half cloud, every cloud is often overseen freely but data and applications are often shared among the mists within the crossbreed cloud [4, 5].

### **5.4 *Community Cloud Model***

Community Cloud is another reasonably distributed computing whereby the arrangement of the cloud is shared physically among numerous associations that have an area with an identical network or zone. The case of such a network is that the place associations/firm's area unit there aboard the money connected foundations/banks [4, 5].

## **6 Analysis of CDM (Cloud Deployment Model)**

Many Organizations need to select a style of Cloud computing relies on their needs and aspirations towards cloud computing. At now, the concept of cloud computing should be thought-about by all businesses. Seeing that it's not solely straightforward however additionally extremely cost-efficient to modify them to perform across the world effortlessly. If you compromise for a server that's public, private, group, or hybrid, the thanks to get into business is cloud computing, however it will solely run if it's vital to your structure wants. Depending on alternative variables, these four preparation models will see a substantial variance. They serve to answer the broad queries on however shared cloud services are often deployed. The expense is sometimes low for the end-user of public clouds, and there aren't any capital investments concerned. Usage of personal clouds needs cost, however thanks to the larger degree of aggregation and resource pooling of personal clouds, the expenditure remains smaller than the expense of owning and running the infrastructure. non-public clouds additionally offer a lot of support for security and compliance than public cloud support.

## **7 Personal (Private), Public and Hybrid Cloud Comparison**

See Table 2.

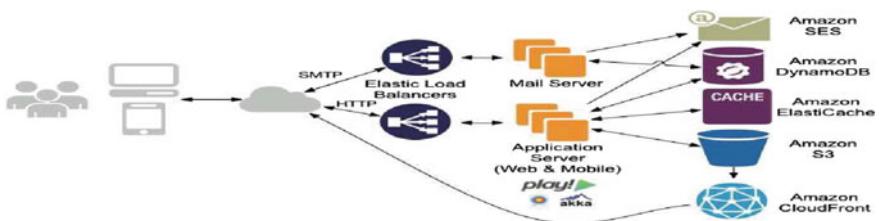
**Table 2** Personal (private), public and hybrid cloud comparison

Attribute	Public cloud	Private cloud	Hybrid cloud	Community cloud
Portability	Easy	Easy	Difficult	Easy
Data security solutions	Better secret writing techniques (authentication of the user by the cloud provider)	Encryption of information on non-public servers	Encryption of private information, sensitive information mustn't be kept publicly cloud	Encryption of private information, sensitive information ought to be kept in community cloud
Performance	Medium	Very good	Good	Very good
Workload	Normal workload	Mission-critical employment with security considerations	Highly dynamic or changeable	High to low
Space required	Low	Large	Medium	Large
Reliability	Medium	High	Medium to high	Very high
Scalability	Very high	Limited	Very high	Very high
Cost of use	Pay-as-use basis	High cost	Pay-as-use basis	Only for community user
Data and application integration	Easy	Easy	Difficult thanks to amendment in cloud platforms	Easy and hard
Security	Low	High	Moderate	Moderate
System management	Easier	Easier	Difficult as compared to public and private cloud	Easy

## 8 Different Cloud Computing Service Providers

### 8.1 AWS

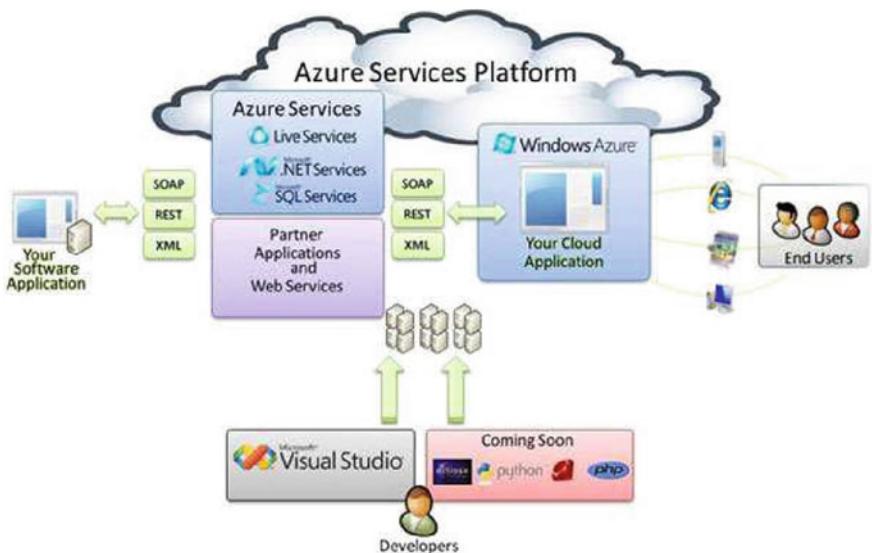
Cloud computing from Amazon, referred to as Amazon net Services (AWS) [3], may be a cloud leader. Yea Computing. AWS may be a complete cloud platform with net services covering the IT market. For instance, servers, storage, databases, messaging, load reconciliation, content infrastructure, computing distribution, and distribution. The inexpensive advantages of Amazon net Services, Elasticity, and sturdiness can assist in organizing. The Amazon net Service design consists primarily of 4 Amazon Elastic work out Cloud (EC2) elements, Amazon fast Storage Facility, (S3), Amazon fast dB, and straightforward Queue Service from Amazon (SQS). the other additional price Service elements that execute plugging that improve attributes and functions are accessible (Fig. 3).



**Fig. 3** Architecture of AWS [6]

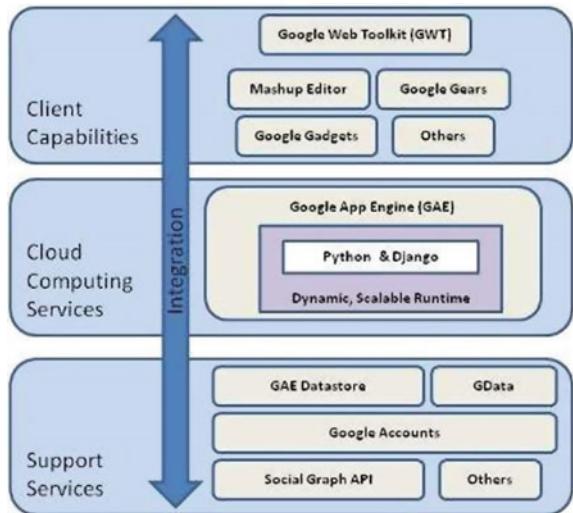
## 8.2 Microsoft Azure

Microsoft Azure may be a cloud computing service developed for the creation, implementation, and management of package and services by Microsoft. Via a world network of knowledge centers operated by Microsoft. It offers package as a service, as a service portal, and Infrastructure It supports many completely different programming languages, tools, and frameworks as a service, together with each Microsoft-specific and Microsoft-specific package and applications from a 3rd party. In Gregorian calendar month 2008, Azure was discovered and discharged as Windows Azure on Gregorian calendar month one, 2010 (Fig. 4).



**Fig. 4** Architecture of Microsoft Azure [6]

**Fig. 5** Architecture of Google Cloud [6]

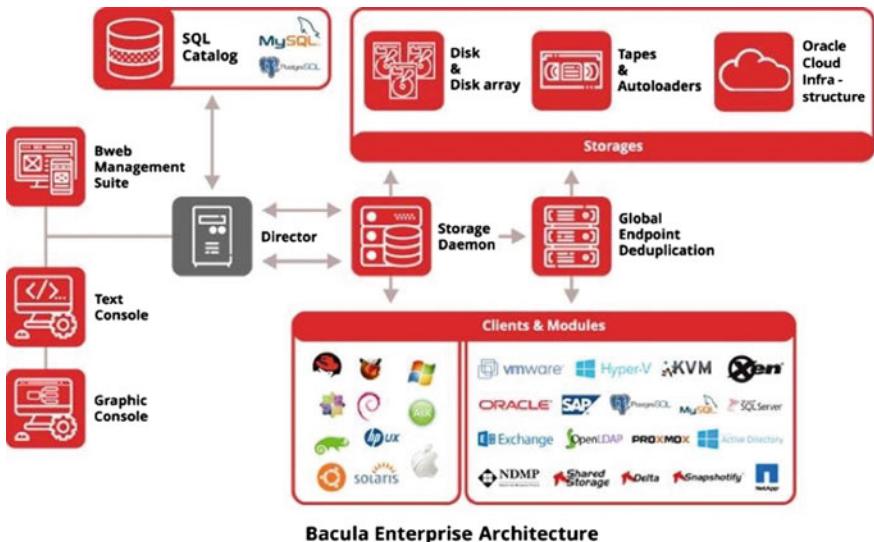


### 8.3 Google Cloud

Google's Google Cloud Platform (GCP) could be a suite of cloud computing services that operates on identical infrastructure used internally by Google for its end-user merchandise, like Google Search, Gmail, file storage, and YouTube. It offers a variety of ascended cloud services, as well as computing, information storage, information analytics, and machine learning, additionally to a group of management tools. The Google App Engine could be a Platform as a Service (PaaS) and is that the internal Google App Engine [7] (Fig. 5).

### 8.4 Oracle Cloud

Oracle Cloud Infrastructure (OCI) is Associate in Nursing IaaS that gives on-site, superior computing capability to work the IT workloads of cloud-based and enterprise corporations. By integration Oracle's autonomous services, advanced security, and serverless computing, OCI offers time period physical property for enterprise applications. Public Cloud on the market. Oracle Cloud may be a cloud computing service provided by Oracle Corporation via a worldwide network of knowledge centers operated by Oracle Corporation that gives servers, storage, networks, applications, and services. The business permits these services to be delivered over the net on demand. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and code as a Service (SaaS) area unit provided by Oracle Cloud (Fig. 6).



**Fig. 6** Architecture of Oracle Cloud [6]

## 8.5 IBM Cloud

IBM Cloud provides a next-generation hybrid cloud design, advanced information and AI technologies, and broad business expertise spanning twenty industries with the foremost open and stable public cloud for business. IBM Cloud may be a series of business cloud computing services provided by IBM, a global IT company. It combines platform as a service (PaaS) with infrastructure as a service (IaaS). The platform scales and serves each giant enterprise companies and tiny development groups and organizations. It's deployed internationally through information centers across the planet (Fig. 7).

## 9 Analysis of CSP

There are various specialist co-ops for distributed computing, for example, Salesforce.com, Microsoft Azure, Google Cloud, and Google Applications Engine, Amazon web Services, IBM cloud arrangements. Among these every one of the similar Analysis of initial three has been done and introduced in the Figure (Table 3):

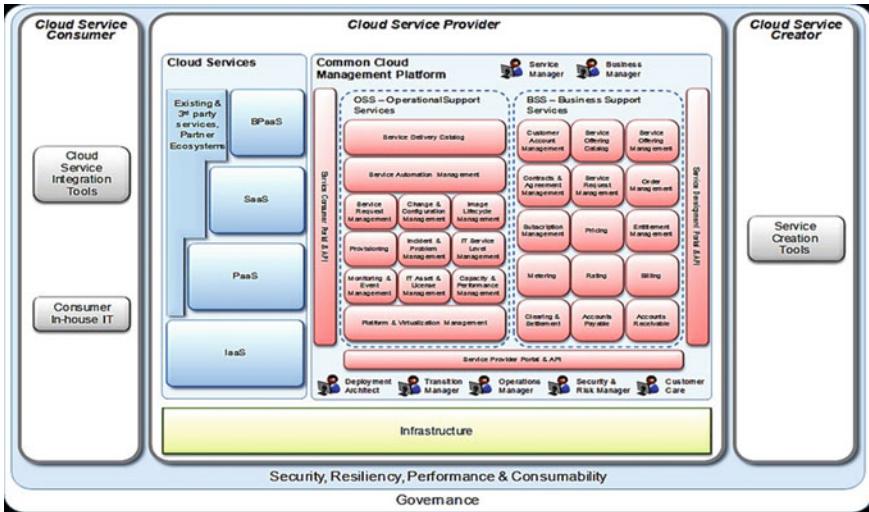


Fig. 7 Architecture of IBM cloud [6]

## 10 Comparing According to Public Cloud Systems

Salesforce.com, Microsoft Azure, Google Cloud, and Google Applications Engine, Amazon web services, IBM cloud arrangements, for example, are different specialist co-ops for distributed computing. Among these, each of the initial three comparable studies according to Public Cloud has been performed and introduced in the figure (Table 4):

## 11 Conclusion

This paper self-addressed the elements and similarities of assorted cloud service suppliers. Every cloud supplier has completely different characteristics between these suppliers, like platform support, thanks to language support, DB support, Autoscaling, load leveling, and fault tolerance, practicableness is outlined during this paper with many elements and sensible extensions to the current paper Microsoft Windows Azure Implementation. Cloud computing plays a really necessary role within the safe and stable provision of resources to be used. The paper compares the Microsoft Services, Google Services, and AWS services. With respect to this analysis, there square measure many potential implications for any analysis. In their square measure, of course, lots of and thousands of suppliers of cloud computing offered to supply services.

**Table 3** Analysis of three major cloud services provider

	Microsoft Azure	Amazon AWS	Google Cloud
Platform supported	Operating systems software	Operating systems software	Runtime software
Cloud services	Infrastructure as a Services Software as a Services	Infrastructure as a Services Platform as a Services	Platform as a Services
Language supported	Python .NET Node.js Power Shell C# F# VB.NET C#.NET PHP BASIC JAVA	Lambda .Net Java Script C++ Java Python Go Ruby PHP	Python, Java, Node.js, Go, Ruby, PHP, or .NET
SLA	99.9% uptime	Amazon S3 available with a monthly uptime 99.9%. Amazon EC2 available with annual uptime 99.5%	100% uptime
Security	Active directory	Identity and Access Management (IAM)	Cloud identity access management
Integrated DB support	Sql Azure	Amazon Aurora, MySQL, PostgreSQL, Elastic cache	It doesn't support any external databases; it provides a knowledge store of its own
Free limit	14,500 Rs. credit in your for use all services	Dynamo 25 GB AWS Lambda 1 Millions, simple notifications Services 1 Million Glaciers 10 GB	<b>300 \$</b> credit in your for use all services
Pricing model	Per minute—rounded up commitment (prepaid and monthly)	Per hours—rounded up (on demand)	Per hours—rounded up (on demand) per 10 min

**Table 4** Comparing according to public cloud systems

Category	Services	Microsoft Azure	Amazon AWS	Google Cloud
Compute	Virtual server	Azure Virtual Machine	EC2 (Elastic Computer Cloud)	Compute Engine
Compute	Container management services	Azure Kubernetes Services (AKS)	Amazon EC2 Container	Kubernetes Engine
Storage	Object storage	Azure Blob Storage	S3 (Simple Storage Services)	Cloud Storage
Storage	Virtual machine disk storage	Azure Page Blobs/Premium Storage	EBS (Elastic Block Storage)	Persistent Disk
Database	Relational database management services	Azure SQL Database	Amazon Aurora and Amazon RDS	Cloud SQL SQL Spanner
Database	Non relational database services	Azure Cosmos DB	Amazon Dynamo DB	Cloud Big Table
Migration services	Database migration	Azure Database Migration Services	AWS Database Migration Services	Not available
Migration services	Server migration	Site Recovery	AWS Server Migration Services	Not available
Networking and content delivery	Virtual networking	Azure VNet	VPC (Virtual Private Cloud)	Cloud Virtual Network
Developer tools	Cloud software development kit	Azure Boards, Azure Pipeline	AWS Cloud9, AWS Code Star	Cloud Source Repositories
Big data and advance analytics	Big data manage cluster as service	Azure HDInsight	Amazon EMR (Elastic Map Reduce)	Cloud Data Proc
Artificial intelligence	Image recognition AI	Emotion API, Face API	Amazon Rekognition	Vision API
Artificial intelligence	Speech recognition AI	Bing Speech API	Amazon Polly	Translation API

(continued)

**Table 4** (continued)

Category	Services	Microsoft Azure	Amazon AWS	Google Cloud
Mobile services	Mobile app development services	Azure Mobile APP	AWS Mobile HUB	Cloud Mobile APP

## References

1. I. Foster, Y. Zho, I. Raicu, S. Lu, Cloud computing and grid computing 360-degree compare, in *Grid Computing Environment's Workshop* (2008), pp. 1–10
2. T. Velte, A. Velte, R.C. Elsenpeter, Cloud computing: a practical approach (chapter 10)
3. H. Yang, M. Tate, A descriptive literature review and classification of cloud computing research. *Commun. Assoc. Inf. Syst.* **31** (2012)
4. G. Garrison, S. Kim, R.L. Wakefield, Success factors for deploying cloud computing. *Commun. ACM* **55**, 62–68 (2012)
5. S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, Cloud computing—the business perspective. *Decis. Support Syst.* **51**, 176–189 (2011)
6. Architecture image from Google
7. P. Peddi, *Design of Simulators for Job Group Resource Allocation Scheduling in Grid and Cloud Computing Environments*, vol. 6, no. 8 (2017), pp. 17805–17811. ISSN: 2319–8753

# Analyzing Frequency Rhythms in Posed and Spontaneous Emotions Using Time-Intensity of EEG Signals



Ritesh Joshi, Pratima Sen, and Maya Ingle

**Abstract** EEG signals are established to be highly effective in spontaneous emotion recognition while subjects are exposed to some stimuli. However, it is posited that the neural signature of posed (volitional/deliberate) emotions are varying from their spontaneous counterpart. In this study, we present subject dependent analysis of spontaneous and posed emotional evocations using normalized time-intensities of higher frequency rhythms. The time-intensities with respect to posed and spontaneous neutral, happy and sad emotions are calculated and topographic distributions are plotted. The topographic time-intensity plots of subjects with respect to posed and spontaneous emotions inferring the prevailing left cerebral activation in posed evocation as compared to coherent right cerebral activation of electrodes in case of spontaneous emotional evocation.

**Keywords** EEG · Time-intensity · Frequency rhythms · Feature normalization · Topographic plot · Posed emotion · Spontaneous emotion

## 1 Introduction

EEG based emotion recognition is a preferred choice as it can't be imitated similar to acoustic signals or facial expressions. It must be useful in real-life human-computer interaction for education, security, behavioral sciences, psychiatry, deception detection, health care etc. This is an interdisciplinary field involving neuroscience, psychology and computer science. In neuroscience the objective is to realize

---

R. Joshi (✉)

Medi-Caps University, Indore, Madhya Pradesh 452001, India

e-mail: [riteshjoshi.indore@gmail.com](mailto:riteshjoshi.indore@gmail.com)

P. Sen

School of Physics, DAVV, Khandwa Road, Indore, Madhya Pradesh 452001, India

M. Ingle

School of Computer Science and IT, DAVV, Khandwa Road, Indore, Madhya Pradesh 452001, India

the neuronal activation and brain processing of emotions. The anterior lobe and amygdala are the two core regions of the brain that are associated with emotion processing. Studies demonstrated that the anterior lobe appears to accumulate more emotive instigation in comparison to other parts of the brain (parietal, occipital and temporal) [1]. Numerous psychophysiological studies showed the resilient association between human emotions and EEG signals. The human cerebral cortex is the region where EEG neuronal electrical activities are clustered across as physiological traces. It is established that the power spectral rhythms (the most all-inclusive features) of EEG can be used for elementary emotion classification [2]. The EEG signals captured from different areas of the brain lobes namely; frontal, parietal, occipital and temporal carry neural traces pertaining to various cognitive functions [3].

Rhythmic and temporal memory characteristics of EEG signals are exercised to recognize emotions based on arousal and valance of Long–Short-Term Memory (LSTM) neural network. Results demonstrate that arousal emotions are well recognized in low rhythms and valance emotions in higher rhythms [4]. Deep belief networks (DBN) trained using differential entropy feature (DE) from multichannel EEG data recognizes positive, neutral and negative emotions. The results show that emotion detection is more coupled with higher bands of EEG data. With the traditional machine learning methods, latent advancement in rhythmic analysis techniques demonstrated its efficacy in EEG based emotion recognition. The performance improvement using multi-band, multiple-feature fusion and multi-channel approaches is attributed to the fusion of information derived from both time and frequency domain [5].

In the past decade, investigations on emotion detection via biological gestures, are studied using various kinds of models including K-nearest neighbor (KNN), Bayesian networks [6, 7], support vector machines (SVM), **Naïve Bayes** (NB), decision trees [8], correlation and regression analysis [9], linear discriminant analysis (LDA) [9], artificial neural network (ANN) [10], and back-propagation networks (BPN) [11] were used to classify the diverse emotions.

The EEG signal of individuals varies in the temporal response according to the inherent differences in nature, gender, ethos, educational upbringing, and living environs. Also, persons may possibly have different behavioral and/or cortical potential responses although observing the same incident. Hence, EEG distributions of each subject are improbable to stake the collective emotional states, though the stimuli are same. Accordingly, a relative response should be used to correlate emotional levels of individuals.

It is pertinent to remind that the EEG signals are engendered due to the movement of ions amid diverse neural points of the brain. Activation of axon is associated with emotion. The neuron activation occurs at various locations in the brain. Activated neuron is polarized and generates a nerve impulse that can be disseminated to other axons, aggravating a drift of information. This activity is probed through the electrodes which represent the intensity and temporal behaviour of cortical potential. The activity varies amid 0 –200  $\mu$ V, and having frequency alternating from 0.3 to 100 Hz. The ensuing signal of an EEG displays crests associated to presence of

electrical commotion, representing an overall spatial site of brain action. The most adopted technique so far to analyse the EEG signal is based on the rhythmic analysis also known as the frequency bands.

Recognizing the fact that the valence emotions affect the higher frequency range e.g. **beta and gamma band**, the present communication aims at examining the role of intensities of EEG signals in quantifying the emotional levels. For this, we propose Analysis of Frequency Rhythms in Some Posed and Spontaneous Emotions (AFRPSE) using the temporal intensity variations of EEG signals. The efficiency of the proposed AFR-PSE is delineated using experimental results. We demonstrate how the neural signatures of posed and spontaneous emotions diverge. In Sect. 2, we focus on the time-intensity based approach for analyzing frequency rhythms in spontaneous and posed emotions. Section 3 illustrates the experimental setup covering selection of subjects, selection of stimuli and electrode placement. Section 4 outlines the activation dynamics with experimental execution. The results based on the time-intensity of EEG signals are presented in Sect. 5 with discussion of results. Finally, in Sect. 6, we conclude with the findings realized in AFR-PSE using time-intensity approach.

## 2 Intensity Based Approach for AFR-PSE

The proposed AFR-PSE primarily emphases on emotion allied time-intensity attribute and explore the activation dynamics in different frequency rhythms. The timeintensity parameter represents the signal intensity over a period of time, the variation of time function. Posed as well as spontaneous neutral, happy and sad emotional states are analyzed using higher frequency rhythms such as, alpha, beta, and gamma respectively. In this work, we explored the time-intensity feature of EEG in analyzing posed and spontaneous emotional exhibits. Specifically, we extracted time-intensity as the complex conjugate of the normalized electrode values in alpha frequency band for neutral emotion and beta, gamma bands for happy and sad emotions respectively. The time-intensity with respect to 14 electrodes is computed using normalized band data values as shown in Eqs. (1) and (2) respectively. The alpha band is analyzed for neutral emotion, beta and gamma bands for happy and sad emotions respectively. Alpha band is observed to be prominent in recognizing neutral emotion using event related spectral perturbation (ERSP) maps based on valance-arousal model [12] while stronger responses are reported in beta and gamma bands for positive (happy), negative (sad) emotions using differential entropy feature [13]. The responses are calculated using following relation:

$$\begin{aligned} \text{band\_data} = & (\text{band\_data} - \min(\text{band\_data})) / (\max(\text{band\_data}) \\ & + \min(\text{band\_data})) \end{aligned} \quad (1)$$

$$\text{band\_intensity} = \text{band\_data} * \text{conj}(\text{band\_data}) \quad (2)$$

### 3 Experimental Setup

Apposite selection of subjects is crucial for construction of dataset. Eight healthy participants were selected for this study. The age of the participants ranges from 18 to 35 years. The consent form is being filled by each subject stating no objection for recording and use of EEG data for aforementioned study. All the selected subjects did not have any past record of psychological illness nor they were under any medication and all had normal to near normal eye vision.

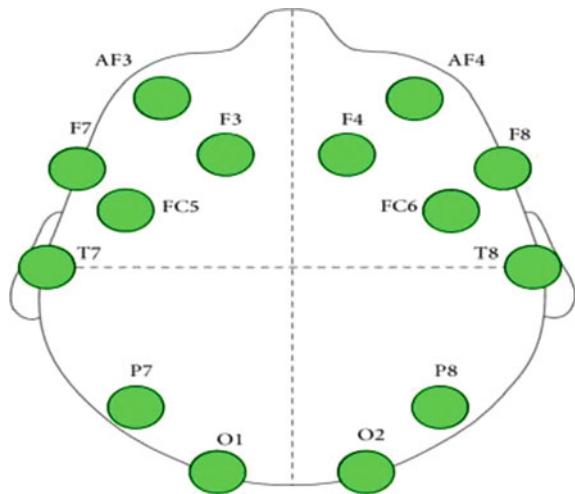
The selection of appropriate stimuli to evoke desired emotion is vital for dataset construction. Short and relevant clips were taken from the movies and soothing videos which depicted happy, sad and neutral emotions. International Affective Picture System (IAPS) image set is standard method to induce emotion based on affective pictures [14]. International Affective Digitized Sound (IADS) database includes acoustic emotional stimuli labelled with arousal and valance values [15]. The set of normative and acoustic emotional stimuli from these databases induce respective emotion in subjects. The coloured images and auditory stimuli are used to elicit emotion. It is posited that audio-visual stimuli educe desired emotion coherently as compared to image or sound stimuli. To induce happy, sad and neutral emotion the subjects are exposed to video clips. Happy emotion was inducted using selected clips from popular Hindi movie “Hera Pheri”, directed by Mr. Priydarshan. For Sad emotion movie clips are selected from popular Hindi movie “Sadma”, directed by Mr. Balu Mahendra and for Neutral emotion clips of soothing music and natural scenes are selected. We made use of video cutter software freely available on internet to get the selected clips [16]. The posed happy, sad and neutral emotions are volitionally evoked by the subjects. The experiment was conceded in the separate room with all experimental setup mounted for spontaneous elicitation and posed exhibition of happy, sad and neutral emotion as shown in Fig. 1.

The 10-20 electrode position system is centered on the relationship between the position of an electrode and the basal region of phrenic layer. The distances between neighboring electrode are either 10 or 20% with respect to the entire anterior-posterior and right-left distance of the cranium [17]. The actual positioning of electrodes as per 10-20 international system is depicted in Fig. 2. Each electrode position has a



**Fig. 1** Experimental setup for data acquisition

**Fig. 2** Placement of electrodes based on international 10-20 system



notelet to categorize the portion and a numeral or another letter to recognize the hemisphere position. The Frontal, Temporal, Central, Parietal and Occipital regions are annotated as F, T, C, P, and O respectively. Numbers 2, 4, 6, 8 and 1, 3, 5, 7 are referring to the right and left hemisphere respectively.

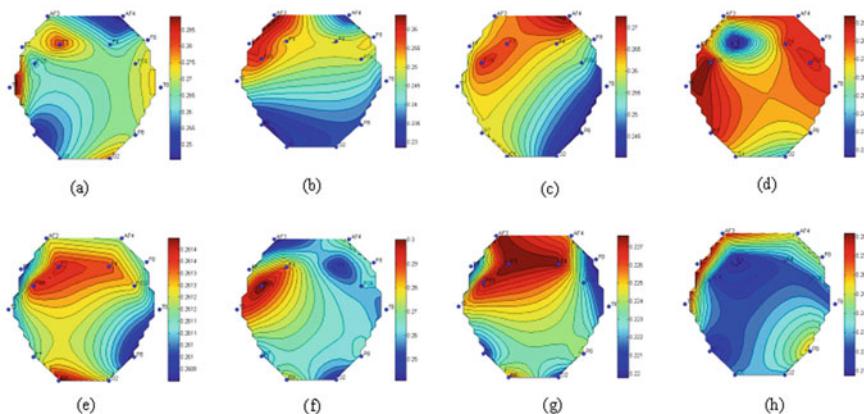
#### 4 Experimental Execution

In order to analyze frequency rhythms in spontaneous and posed happy, sad and neutral emotions, a trial is premeditated to record raw EEG data. The experiment is carried out on Windows 10 using EEGLAB toolbox on MATLAB 13A and time-frequency analysis is performed using Topographic Maps. The 14 channel Emotiv Epoc+ device is used in this study to record raw EEG signals. The sample rate per unit time of device is 128 instances per second and the recorded signals are in European Data Format (.edf). A special experiment was designed to capture the raw EEG data required for AFR-PSE analysis. The raw EEG data is comprises of signals captured from each subject with respect to spontaneous and posed happy, sad and neutral emotion.

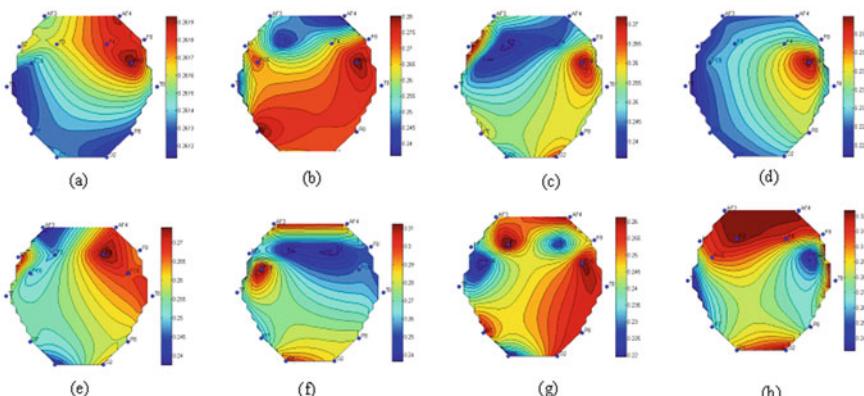
The captured EEG signal is tainted with noise and artifacts and hence needed preprocessing. However, the comprehensive exclusion of noise and artifacts is not feasible with the existing methods and using present available data capturing process. The band pass filtering is used for removing artifacts and noise using EEGLAB [18]. These filtered signals were converted into numerical values for further processing corresponding to spontaneous and posed happy, sad and neutral emotion. The beta (13–30 Hz) and gamma (30–50 Hz) bands are extorted for analyzing happy and sad emotions. The alpha band (8–13 Hz) is extracted from raw EEG signal for analysis

of neutral emotion. The feature vector is constructed for each of the 8 participants based on the filtered dataset using the time-intensity feature extraction phase of the AFRPSE.

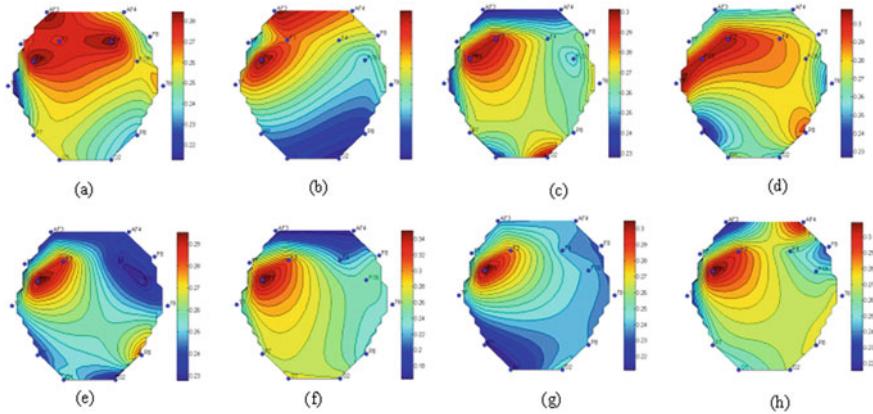
The recording of 60 s is used for analysis with respect to each emotion. As the sampling rate of the device is 128 samples/s and hence 7680 data points are used for analysis. The experiment was carried out on eight subjects varying in the age group of 19–45 years. The topographic distribution of the alpha band obtained at different electrode positions have been plotted for posed and spontaneous neutral emotion as shown in Figs. 3 and 4. The topographic distribution of the beta and gamma band obtained at different electrode positions have been plotted for posed and spontaneous happy and sad emotions as shown in Figs. 5, 6, 7, 8 and 9, 10, 11, 12 respectively.



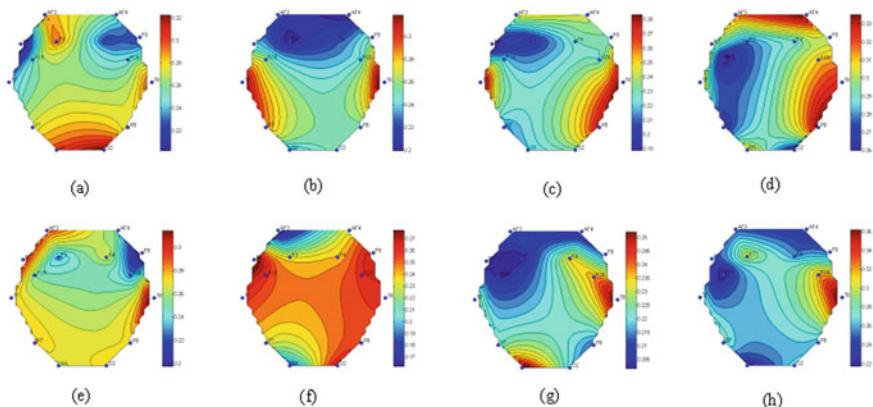
**Fig. 3 a–h** Topographic plot of alpha band for subject 1–8 with respect to posed neutral emotion



**Fig. 4 a–h** Topographic plot of alpha band for subject 1–8 with respect to spontaneous neutral emotion



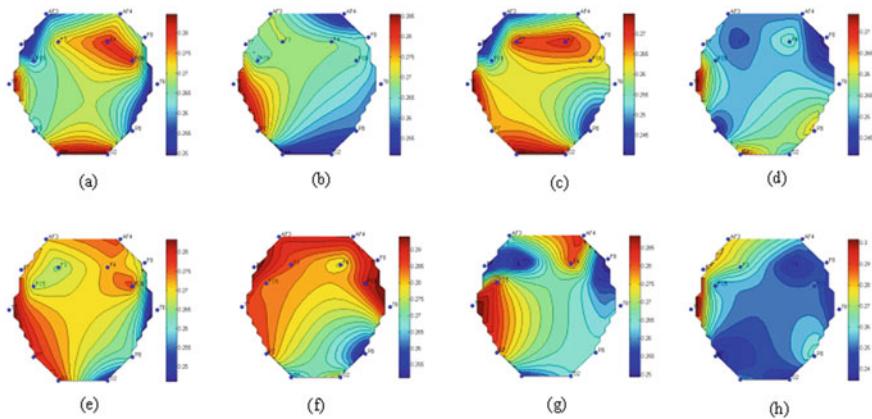
**Fig. 5 a–h** Topographic plot of beta band for subject 1–8 with respect to posed happy emotion



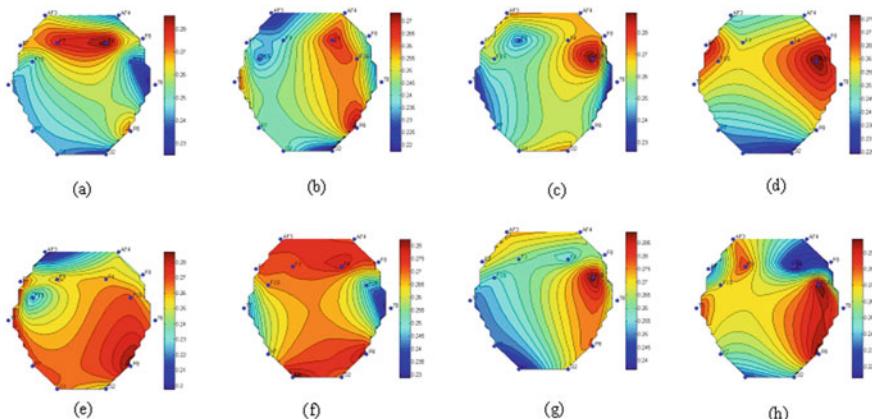
**Fig. 6 a–h** Topographic plot of beta band for subject 1–8 with respect to spontaneous happy emotion

## 5 Results and Discussion

Using AFR-PSE, we have endeavored to investigate the frequency rhythms in posed and spontaneous neutral, happy and sad emotions. Further, subject dependent time-intensity based topographic distributions are plotted for posed and spontaneous neutral, happy and sad emotions respectively. Three cases namely; Case I, Case II and Case III based on three aforementioned emotions have been framed. Here, we discuss observations drawn for aforesaid three cases as below:



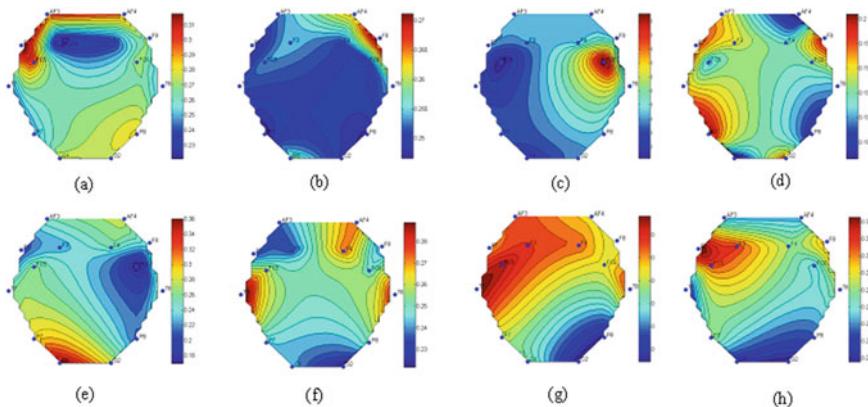
**Fig. 7 a–h** Topographic plot of gamma band for subject 1–8 with respect to posed happy emotion



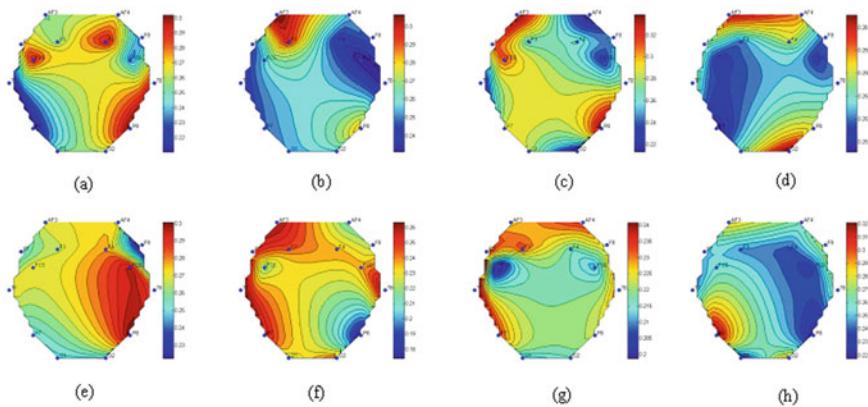
**Fig. 8 a–h** Topographic plot of gamma band for subject 1–8 with respect to spontaneous happy emotion

### Case I: Analysis of Neutral Emotion

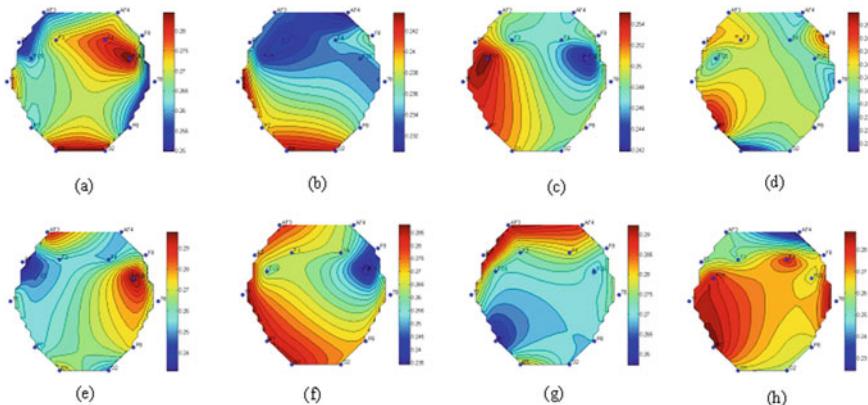
A broad view of topographic maps of subjects indicates that the signals of higher intensity are obtained for posed neutral emotion through the left hemisphere of the brain as indicated by the electrode FC5. Similarly, topographic maps of subjects for spontaneous neutral emotion indicate that the signals of higher intensity are obtained in the right hemisphere of the brain as indicated by the electrode FC6. The findings suggest that posed neutral emotion is predominantly expressed by left hemisphere of the brain, whereas spontaneous neutral emotion primarily concentrated in right hemisphere. However, in case of subject 6 (a stage actor) in both posed and spontaneous neutral evocation left part of the brain is activated, indicating the controlled evocation of neutral emotion. Although the intensity levels as shown in the Figures



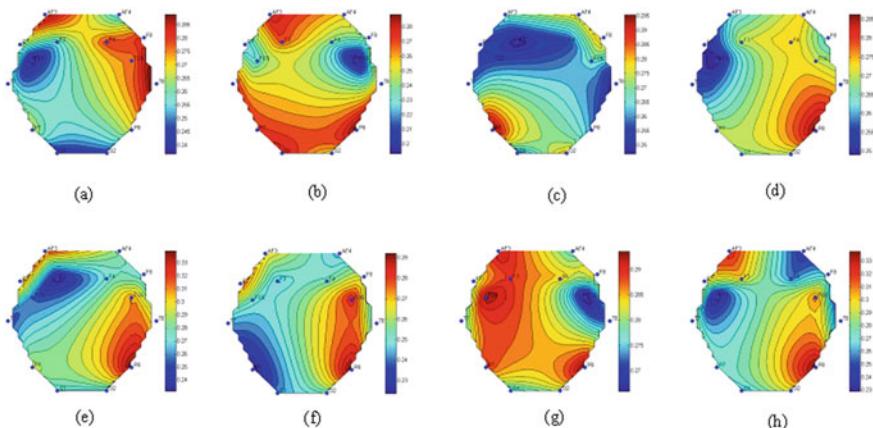
**Fig. 9** a–h Topographic plot of beta band for subject 1–8 with respect to posed sad emotion



**Fig. 10** a–h Topographic plot of beta band for subject 1–8 with respect to spontaneous sad emotion



**Fig. 11** a–h Topographic plot of gamma band for subject 1–8 with respect to posed sad emotion



**Fig. 12 a–h** Topographic plot of gamma band for subject 1–8 with respect to spontaneous sad emotion

are normalized by the maximum intensity obtained by us however, the raw data was found to be higher in case of the younger subject aging between 19 and 23 years.

### Case II: Analysis of Happy Emotion

The comprehensive overview of topographic maps of the subjects drawn using beta and gamma frequency bands for posed happy emotion indicates that the higher intensities are gained in the left hemisphere of the brain as shown by the electrodes FC5 (in case of beta) and T7 (in case of gamma) respectively. On the contrary, topographic maps of subjects for spontaneous happy emotion using beta and gamma bands divulge that higher intensities are predominantly obtained in the right hemisphere of the brain as shown by electrode T8 (in case of beta) and P8 (in case of gamma) respectively for spontaneous happy emotion. However, in case of subject 6 (a stage actor) in both posed and spontaneous happy evocation right part of the brain is activated, indicating the controlled evocation of happy emotion in gamma band. Although the intensity levels as shown in the Figures are normalized by the maximum intensity obtained by us however, the raw data was found to be higher in case of the younger subject aging between 19 and 23 years for happy emotion.

### Case III: Analysis of Sad Emotion

The compendious view of topographic maps of the subjects drawn using beta and gamma frequency bands for posed sad emotion indicates that the higher intensities are gained randomly in the left and right hemisphere of the brain as shown in Figs. 9 and 11 respectively. On the contrary, topographic maps of subjects for spontaneous sad emotion using beta and gamma bands divulge that higher intensities are predominantly obtained in the frontal left hemisphere of the brain as shown by electrode AF3 and parietal right hemisphere electrode P8 respectively. Although the intensity levels as shown in the Figures are normalized by the maximum intensity obtained

by us however, the raw data was found to be higher in case of the younger subject aging between 19 and 23 years.

## 6 Conclusion

The subjects are persuaded with audiovisual stimulus to elicit spontaneous emotions. In posed exhibition, subjects are directed to enact emotions intentionally. The time-intensities of various electrodes for spontaneous emotions are observed to be showing higher activation as compared to posed emotional elicitation. In spontaneous elucidation, subjects experienced respective emotions naturally when subjected to stimulus as compared to posed exposition. Further, time-intensity plot of spontaneous and posed emotions suggesting the comprehensible neural activation of right hemispheric probes in case of natural emotion as compared to predominant left hemispheric activation in posed exhibition. Further, analyzed frequency rhythms in various frequency bands for happy, sad and neutral emotion will assist investigators to distinguish posed and spontaneous emotions with time-intensity features of EEG.

## References

1. N.S. Suhami, J. Mountstephens, J. Teo, EEG-based emotion recognition a state-of-the-art review of current trends and opportunities. *Computational Intelligence Neuroscience*, **16**, 20208875426. <http://doi.org/10.1155/2020/8875426>. PMID 33014031; PMCID PMC7516734 (2020)
2. R. Sarno, M.N. Munawar, B.T. Nugraha, Real-time electroencephalography-based emotion recognition system. *Int. Rev. Comput. Softw. (IRECOS)* **11**(5), 456–465 (2016)
3. J. Yan, S. Chen, S. Deng, A EEG-based emotion recognition model with rhythm and time characteristics. *Brain Inf.* **67** (2019). <http://doi.org/10.1186/s40708-019-0100-y>
4. W.L. Zheng, B.L. Lu, Investigating critical frequency bands and channels for EEG based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Dev.* **7**(3), 1 (2015)
5. C. Liu, P. Rani, N. Sarkar, An empirical study of machine learning techniques for affect recognition in human-robot interaction, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS, Sendai, Japan (2005)
6. G. Rigas, C.D. Katsis, G. Ganiatsas, D.I. Fotiadis, *A User Independent, Biosignal Based, Emotion Recognition Method*. Springer, Berlin, Germany (2007)
7. C. Zong, M. Chetouani, Hilbert-Huang transform based physiological signals analysis for emotion recognition, in *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology, ISSPIT*, Ajman, UAE (2009), pp. 334–339
8. L. Li, J.H. Chen, Emotion recognition using physiological signals from multiple subjects, in *Proceedings of the International Conference on Intelligent Information Hiding and Multimedia*, Pasadena, CA, USA (2006), pp. 437–446
9. F. Nasoz, K. Alvarez, C.L. Lisetti, N. Finkelstein, Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cogn. Technol. Work* **6**(1), 4–14 (2004)
10. A. Haag, S. Goronzy, P. Schaich, J. Williams, Emotion recognition using bio-sensors first steps towards an automatic system. *Lecture Notes in Computer Science*. Springer, Berlin, Germany (2004)

11. J. Kim, E. Andre, Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(12), 2067–2083 (2008)
12. R. Du, R.M. Mehmood, H.J. Lee, Alpha activity during emotional experience revealed by ERSP. *J. Internet Technol.* **15**(5), 775–782 (2014)
13. P. Zhou, J. Wei, S. Liu, J. Meng, X. An, H. Qi, F. He, X. Wang, M. Scott, D. Ming, M. Liu, The intensity recognition of tension emotion based on multi-physiological signals. *Journal of Psychiatry and Brain Science (JPBS)* **5**(1), 1–6 (2016). <https://doi.org/10.20900/jpbs.20160019>
14. P.J. Lang, M.M. Bradley, B.N. Cuthbert, *International Affective Picture System (IAPS) Technical Manual and Affective Ratings*. NIMH Center for the Study of Emotion and Attention, Gainesville, FL (1997)
15. M. Bradley, P. Lang, The international affective digitized sound (IADS-2) affective rating of sounds and instruction manual. Technical Report, University of Florida (2007)
16. <http://www.freevideocutter.com>
17. H. Jasper, The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* **39**, 371–375 (1958)
18. A. Delaureme, S. Makeig, EEGLAB an open source toolbox for analysis of single trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**(1), 9–21 (2004)

# Improving Classification Performance Using the Semi-pivoted QR Approximation Algorithm



Marina Popolizio, Alberto Amato, Vincenzo Piuri, and Vincenzo Di Lecce

**Abstract** Aim of this paper is to present a method to improve the classification performance of a Fuzzy C-means based classifier. The obtained results show that this method can improve the performance of the classifier both in terms of computational efficiency (by reducing the amount of data to be analyzed) and in terms of classification error rate. The proposed method is based on the Semi-Pivoted QR approximation (SPQR) algorithm. It reduces a numeric dataset (a matrix) to its more important features (where each feature is a column of the matrix). The framework discussed in this article can be used by researchers and practitioners to set up high-performance machine learning systems.

**Keywords** Semi-pivoted QR approximation (SPQR) · Fuzzy C-means classifier · Data quality

## 1 Introduction

In the latest years, expert systems based on machine learning and other advanced techniques are becoming ever more common in everyday life. Undoubtedly, the quality and reliability offered by these systems is highly dependent on the quality and quantity of the data that they are using. The higher the quality of the input data, the better the results obtained by machine learning systems built on the data. However, the qualities of the data are hidden, so it is not easy to check if the data has

---

M. Popolizio (✉) · A. Amato · V. Di Lecce

Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari, Via Orabona 4, Bari, Italy

e-mail: [marina.popolizio@poliba.it](mailto:marina.popolizio@poliba.it)

V. Di Lecce

e-mail: [v.dilecce@poliba.it](mailto:v.dilecce@poliba.it)

V. Piuri

Dipartimento di Informatica, Universita' degli Studi di Milano, Via Celoria 18, Milano, Italy

properties suitable for use in machine learning. The risk is common to both small datasets (often used in medicine) and big datasets (often environmental).

In this paper, the attention is focused on environmental data collections with large amounts of time-driven data. The data evaluation process, performance improvement strategy, and data quality assessment framework discussed in this article can be used by researchers and machine learning practitioners to set up high-performance machine learning systems.

Aim of data quality evaluation is to obtain a exhaustive characterisation and measurement of data, both for quantitative and qualitative properties. In [1, 2], authors defined data quality as “data suitable for use by data consumers” but also as “the quality of the mapping between a state of the real world and a state of the information system”.

According to [3], the three most important quality features for machine learning performance are completeness, correctness and variety. “Data quality” evaluation could concern measuring wrong or missing data in the overall context [4]. In general, data quality means meeting the needs and preferences of its users or tasks or, more succinctly, the ability of the data to be fit for purpose. In our case, one of the aims of this study is to understand in advance whether a dataset can be used to build a high-performance machine learning system [5, 6]; selected quality dimensions, such as completeness, correctness and variety, will be used to assess the quality of the data [1].

Quality is defined by Ref. [7] as “the degree to which a set of intrinsic characteristics of an object satisfies requirements” or, according to [8] as “the degree or level of confidence that the data provided satisfies the requirements of the user”. This definition is not easily applicable to datasets although it must be kept in mind. There are many approaches to data quality assessment [9–11] and almost all of them are directly related to the construction of machine learning systems. This activity is however important because it can heavily influence the performance of a machine learning system built on data.

Data quality must be evaluated systematically to ensure the quality of the application, especially in deep learning.

We will not go into the specifics of the definitions because they are already referred to in the bibliography. Relevant is also the discussions about noise. The results reported in [3] shown that data with noise [12] could still be effective for training a deep learning model. But, in our opinion, the performance of deep learning systems can be improved by removing noise from that dataset used to train them. In [13] there is a discussion about the impact of the size and quality of a dataset on the performance of machine learning also in the case of deep learning.

In this work the authors propose a method to improve the performance of a fuzzy clustering classifier applied to an environmental dataset using the SPQR algorithm to select the features of the dataset that give a better classification performance.

The remaining part of the paper is so organized: Sect. 2 proposes a brief discussion about the Fuzzy C-Means algorithm while Sect. 3 describes the cluster silhouette parameter used to evaluate the clustering quality. Section 4 introduces the SPQR

algorithm; Sect. 5 reports some numerical experiments, final remarks and conclusions are reported in Sect. 6.

## 2 Fuzzy Clustering and Classifier

Clustering is an unsupervised learning technique used to group objects according to some similarity inherent among them. In literature various methods for clustering the objects are discussed as, for example, hierarchical, partitional, grid, density and model based [14].

In this work, the Fuzzy C-Means (FCM) algorithm has been used. It is a well-established clustering algorithm based on the optimization of the objective function

$$Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|x_k - v_i\| \quad (1)$$

where:

- $\|\cdot\|$  denotes a distance function (i.e. Euclidean, Mahalanobis, etc.)
- $v_1, v_2, \dots, v_c$  are the prototypes (centroids) of the constructed clusters
- $X = \{x_1, x_2, \dots, x_N\}$  is the set of the points to be clustered
- $U = [u_{ik}]$  is the partition matrix
- $c$  is the number of clusters
- $N$  is the number of points to be clustered
- $i = 1, 2, \dots, c$
- $k = 1, 2, \dots, N$
- “ $m$ ” is a fuzzification coefficient greater than 1 whose function is to control a level of fuzziness of the clusters (their membership functions).

The minimization of  $Q$ , that is carried out with respect to the partition matrix and the prototypes, gives rise to the structure in  $X$ . The generic optimization scheme is an iterative procedure where the values of the partition matrix and the prototypes are successively updated. The process terminates when some stopping criterion has been met. For example, if  $U$  and  $U'$  denote partition matrices obtained in two consecutive iterations, the procedure may stop when the quantity gets smaller than some predefined positive threshold.

$$\|U - U'\| = \max_{i,k} |u_{ik} - u'_{ik}| \quad (2)$$

It should be noticed that the discovery of the structure in the dataset  $X$  is completed in a fully unsupervised mode. The objective function somehow reflects the geometry of the data so that the minimization of  $Q$  leads to an “acceptable” structure.

Starting from these considerations, the hidden structure revealed by a clustering algorithm can be used as a classifier applying the following rules:

- The “anchor points” of the classifier are the prototypes of the clusters
- Each cluster defines a class
- A point  $\mathbf{x}$  belongs to a class defined by the cluster with prototype  $\mathbf{v}_j$  if:

$$j = \arg(\min_i ||\mathbf{x} - \mathbf{v}_i||^2) \quad (3)$$

### 3 Silhouette Method for Clustering Evaluation

In order to evaluate the performance of a clustering method it is necessary to compare its results to either standard results, or to the results of another method. Dealing with problems where the classification of each point is known a priori, it is possible to use this knowledge to evaluate the classification results obtained by the clustering algorithm. Indeed, in these cases, it is possible to count the number of misclassified patterns to compute the classification error rate. By the other hand, there are problems where there is no a priori knowledge about the dataset under analysis and so it is not possible to evaluate the clustering performance using the previous method. To face this kind of problem, various solutions have been proposed as shown in [15]. Furthermore, in the latest years, the silhouette parameter is becoming a more commonly used method for assessing the quality of clusters [16].

This method proposes a comparative measure of the similarity of an object to its cluster (tightness) and its similarity to the other clusters (separation). This parameter is defined as follows: if  $x$  is a point belonging to cluster **A**, then  $a(x)$  is the mean distance between  $x$  and all other points of **A**. Let us now consider any cluster **B** different from **A** and compute the average distance between  $x$  and all points of **B** ( $d(x, B)$ ). Once we have computed  $d(x, B)$  for each cluster **B** such that **B**  $\neq$  **A**, we select the smallest of these numbers and denote it by

$$b(x) = \min_{A \neq B} d(x, B) \quad (4)$$

Starting from these considerations, the silhouette for the point  $x$  is defined as shown in the following formula:

$$(x) = \begin{cases} 1 - \frac{a(x)}{b(x)} & \text{if } a(x) < b(x) \\ 0 & \text{if } a(x) = b(x) \\ \frac{b(x)}{a(x)} - 1 & \text{if } a(x) > b(x) \end{cases} \quad (5)$$

From this definition it is possible to say that for each point  $x$  into the dataset:

$$-1 < s(x) < +1$$

An in-depth analysis of this parameter is reported in [16]. Furthermore, this method proposes an effective graphical representation of how well each object has

been classified. Any distance metric can be used to compute the silhouette such as the Euclidean or the Manhattan distance.

Due to all these characteristics, silhouette is often used to evaluate the performance of clustering algorithms when they are applied to datasets for which there is no a-priori knowledge (that is also the case study of this paper). Furthermore, the optimal number of clusters of a given dataset may be chosen through the silhouette performing an iterative process:

1. compute the clustering by setting a certain number  $C$  of clusters
2. computing the silhouette for each point into the dataset
3. if there is a good percentage of points with high silhouette value then  $C$  is a good value for number of clusters, otherwise increase the value of  $C$  and repeat the process of clustering.

## 4 SPQR Algorithm

In data applications, a matrix  $A$  composed of  $m$  row and  $n$  columns is used to represent  $m$  objects characterized by  $n$  features. So, this matrix has dimension  $m \times n$  and  $m$  is usually very large when dealing with high-dimensional data, such as images or environmental datasets.

This large dimension gives rise to some troubles, like the difficulty in visualizing the data and often in their storage. However, some of the represented features are often irrelevant and/or redundant and, when their number is remarkable, they can significantly affect the analysis of the data. *Feature selection* is the process of selecting the most meaningful features thus to reduce the dimensionality of the data, to facilitate their visualization and understanding and, in some cases, it works as preprocessing for tasks like classification, regression or clustering. Due to the importance of this phase, in the last decades, numerous feature selection methods have been introduced in different areas such as pattern recognition, machine learning, data mining and statistical analysis. Feature selection is also useful for reducing training times, avoiding overfitting, and facilitating data understanding. Many of these methods aim to approximate the matrix  $A$  by means of a “smaller” matrix obtained by combining its columns and rows. The drawback of these methods is that they usually yield dense factorizations and, more seriously, these terms are often much harder to interpret than the original ones. For example, truncating the SVD at  $k$  terms is one of the most common way to get the “best” rank- $k$  approximation to  $A$  when measured with respect to any unitarily invariant matrix norm. Unfortunately, it is well known that this representation is difficult to relate with the underlying data and processes generating that data. The PCA method presents the same disadvantage. For these reasons methods which aim to use just a small part of the original matrix are of particular interest. Specifically, many methods focus on the “column subset selection problem”, that is to say, their goal is to find a subset of  $k$  actual columns of  $A$ , with  $k$  much smaller than  $n$ , which “captures” most of the information of  $A$ , with respect to the spectral or the Frobenius norm.

Essentially two classes of methods may be defined: *randomized* methods, which use probability distributions to find the most representative columns in a matrix, and the *deterministic* methods, which apply a deterministic procedure to select and return exactly  $k$  columns.

An effective deterministic method to reduce the matrix  $A$  to its more important columns is the “Semi-Pivoted QR approximation” (SPQR) due to Stewart [17–19]. As its name suggests, the key approach is the QR decomposition that expresses  $A$  as the product of an orthogonal matrix  $Q$  and an upper triangular matrix  $R$ . These factors are computed by the *Gram–Schmidt* algorithm for orthonormalizing the columns of  $A$  one at a time, from first to last. In many situations the *Pivoted* QR has to be preferred: it differs in that the Gram–Schmidt procedure takes the largest remaining column at the beginning of each new step and thus introduces a permutation matrix  $P$  such that

$$A \cdot P = Q \cdot R \quad (6)$$

When  $A$  is rank deficient the column pivoting  $A \cdot P$  is applied to improve the numerical accuracy. Moreover,  $P$  is usually chosen so that the diagonal elements of  $R$  are nonincreasing, a specific feature which is useful in the following. More in detail, we may partition the expression above as

$$\begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \quad (7)$$

and the following properties hold:

1.  $B_1 = Q_1 \cdot R_{11}$
2.  $\|B_2 - Q_1 \cdot R_{12}\| = \|R_{22}\|$

The *semi-QR* algorithm exploits these results to use the approximation

$$A \cdot P \approx Q_1 \cdot [R_{11} \ R_{12}] \quad (8)$$

that, thanks to the property 1 above, reproduces the first  $k$  columns of  $A \cdot P$  exactly, by introducing a quantifiable error (property 2). An additional strength of this method is that the explicit computation of the nonsparse orthogonal matrix  $Q_1$  is not required. In practice, given a rank parameter  $k$ , the SPQR algorithm gives the  $k$  columns of  $A$  whose span approximates the column space of  $A$ ; they form the matrix  $B_1$  of dimension  $m \times k$  while the nonsingular upper-triangular  $k \times k$  matrix  $R_{11}$  orthogonalizes these columns.

## 5 Experiments and Results

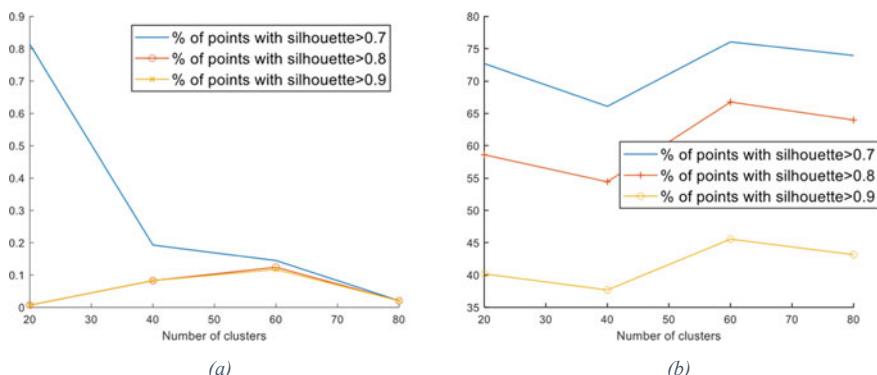
The SPQR algorithm has been applied to an environmental dataset containing meteorological and geomorphological data about various climatic events occurred in Apulia region in Italy in the recent years. The dataset is composed by 14,502 events each one described using 25 parameters. These parameters are geomorphological characteristics of the area interested by the event and some meteorological data referred to the 24 h before the peak of the event (temperature, rain intensity in 1 h, 3 h, 6 h, 12 h and 24 h, wind speeds and directions).

A fuzzy C-means algorithm has been used to cluster this dataset. Since there is no apriori knowledge about the number of the classes into the dataset, the iterative process described in Sect. 3 has been applied to find the better number of clusters using the silhouette parameter as criterion of evaluation. All the experiments have been carried out using Matlab.

Figure 1a the percentage of points with a silhouette respectively greater than 0.7, 0.8 and 0.9 is shown as a function of the number of clusters. The number of clusters varies from 20 to 80.

After this stage, the SPQR algorithm has been applied to the dataset to select the most representative features (columns) of the dataset. From this elaborated dataset, the first three columns have been extracted to be used in the successive clustering analysis. The same process described above has been applied to the reduced dataset obtaining the results shown in Fig. 1b.

From this comparative evaluation it emerges that the FCM works better for the reduced dataset than for the original one. Indeed, for the former the results in terms of patterns correctly classified in terms of silhouette are better than those obtained for the latter. Furthermore, the proposed method gets a strong speed up into the classification process, since it reduces the quantity of data to be analyzed by the clustering algorithm.



**Fig. 1 a, b** Percentage of points with silhouette  $>0.7$ ,  $>0.8$  and  $>0.9$  as the number of clusters varies. The algorithm has been applied to the original dataset (a) and to the reduced dataset (b)

## 6 Conclusions

Thanks to the latest technological improvements, in recent years the diffusion of complex applications based on the analysis of large amounts of data has spread. Data mining algorithms, expert systems and automatic classifiers are some examples of such applications. The results that these systems give are strongly influenced by the quality of the input data. Aim of data quality evaluation is to obtain a comprehensive characterization and measurement of quantitative and qualitative properties of data. The authors follow the definition of “data quality” proposed by Refs. [1, 2] that considers data quality as “data suitable for use by data consumers”. Starting from this consideration, clustering silhouette has been used as a parameter to assess the performance of the clustering algorithm.

Fig. 1a shows the obtained results on a dataset containing environmental data. These results have been compared to those obtained applying the same clustering algorithm to a reduced dataset obtained from the original one using the SPQR algorithm.

Fig. 1b shows the obtained results.

This comparative evaluation shows that using the SPQR algorithm improves the quality of input data obtaining better results in complex applications such as automatic classifier: the enhancement is both in terms of computing efficiency, since the algorithm works on less data, and in terms of classification error.

## References

1. R.Y. Wang, D.M. Strong, Beyond accuracy: what data quality means to data consumers. *J. Manage. Inf. Syst.* **12**(4), 5–33 (1996)
2. Y. Wand, R.Y. Wang, Anchoring data quality dimensions in ontological foundations. *Commun. ACM* Vol. **39**, N. 11, pp. 86–95, (1996)
3. J. Ding, X. Li, X. Kang, V.N. Gudivada, A case study of the augmentation and evaluation of training data for deep learning. *J. Data Inf. Qual.* **11**(4), 1–22 (2019)
4. M. Popolizio, A. Amato, F. Liquori, T. Politi, A. Quarto, V. Di Lecce, The GAIN method for the completion of multidimensional numerical series of meteorological data. *IAENG Int. J. Comput. Sci.* **48**(3), 496–506 (2021)
5. H. Chen, G. Cao, J. Chen, J. Ding, A practical framework for evaluating the quality of knowledge graph, in *Proceedings of China Conference on Knowledge Graph and Semantic Computing* (2019), pp. 111–122
6. J. R. Evans, W.M. Lindsay, *The Management and Control of Quality*, vol. 5 (South-Western, Cincinnati, OH, USA, 2002), pp. 115–128
7. Quality Management Systems—Fundamentals and Vocabulary (ISO 9000:2015), ISO, NF EN ISO 9000
8. Standards for Processing Aeronautical Data, RTCA, Inc., DO200B (June 2015)
9. C. Batini, C. Cappiello, C. Francalanci, A. Maurino, Methodologies for data quality assessment and improvement. *ACM Comput. Surv.* **41**(3), 1–52 (2009)
10. Chen et al., Data evaluation and enhancement for quality improvement of machine learning. *IEEE Trans. Reliab.* (IF4.424) (2021). <http://doi.org/10.1109/tr.2021.3070863>

11. F. Daniel, P. Kucherbaev, C. Cappiello, B. Benatallah, M. Allahbakhsh, Quality control in crowdsourcing: a survey of quality attributes, assessment techniques, and assurance actions. *ACM Comput. Surv.* **51**(1), 1–40 (2018)
12. J. Krause et al., The unreasonable effectiveness of noisy data for finegrained recognition, in *Proceedings of European Conference on Computer Vision* (2016), pp. 301–320
13. C. Sun, A. Shrivastava, S. Singh, A. Gupta, Revisiting unreasonable effectiveness of data in deep learning era, in *Proceedings of IEEE International Conference on Computer Vision* (2017), pp. 843–852
14. A. Saxena, M. Prasad, A. Gupta, N. Bharill, O.P. Patel, A. Tiwari, M. Er, C.-T. Lin, A review of clustering techniques and developments. *Neurocomputing* **267** (2017). <http://doi.org/10.1016/j.neucom.2017.06.053>
15. W.M. Rand, Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **66**(336), 846–850 (1971). <https://doi.org/10.2307/2284239>
16. P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 5365 (1987). ISSN 0377-0427
17. M.W. Berry, S.A. Pulatova, G.W. Stewart, Computing sparse reduced-rank approximations to sparse matrices. *ACM Trans. Math. Softw.* **31**, 252–269 (2005)
18. G.W. Stewart, Four algorithms for the efficient computation of truncated pivoted QR approximations to a sparse matrix. *Numer. Math.* **83**, 313–323 (1999)
19. G.W. Stewart, Error analysis of the quasi-Gram–Schmidt algorithm. *SIAM J. Matrix Anal. Appl.* **27**(2), 493–506 (2004)

# Comparative Study on Video Based Face Recognition Methods



Heta S. Desai and Atul M. Gonsai

**Abstract** Face recognition has received tremendous support in the discipline of computer, which is used for automatic attendance from CCTV (closed circuit television) cameras, monitoring criminal activity, and person tracking systems, etc. Tracking through video surveillances includes pre-processing, feature extraction, face detection and face recognition. Many problems are there when detecting faces during process of face recognition through surveillance camera like low quality camera, illumination problem, variation in face angle, occlusion problem, variation in facial expression, etc. some of these are still required more attention which is illumination and occlusion. When human faces are covered with some objects like face mask, sunglasses, beard, scarf and many other than face recognition accuracy have been drastically decreased. As many researchers have worked on tracking individuals from Surveillance camera, which found that still occlusion along with illumination is major problem when CCTV camera recognizes the face from the live feed as specially faces with the face mask. There are numerous algorithms and classifiers that can identify and comprehend faces of person from the front as well as with certain changes in an orientation of the face, but not many researches are utilized for unconstrained and occluded with illumination video-based face Recognition. So, this research review concentrates on various diverse algorithm for face recognition with occlusion and illumination from the surveillance camera and its parameter like execution time as well as recognition accuracy.

**Keywords** Face detection · Face mask · Face recognition · Occlusion · Person tracking · Surveillances camera

---

H. S. Desai (✉) · A. M. Gonsai  
Saurashtra University, Rajkot, India  
e-mail: [prof.hetadesai@gmail.com](mailto:prof.hetadesai@gmail.com)

A. M. Gonsai  
e-mail: [amg@sauuni.ac.in](mailto:amg@sauuni.ac.in)

## 1 Introduction

In recent years, foremost emerging research technologies in the fields of computer discipline is pattern matching, pattern recognition, biometrics, face detection and recognition has received special attention. Face detection and face recognition have been studied a lot because it helps in security, crime detection, missing persons tracking, automatic university attendance, etc. Nowadays, video-based face recognition has become a vibrant research topic due to its huge variety of uses in residential area as well as crowded area.

Surveillance is a one form of approach and it moreover guiding to keep track of the object with the resource of the use of recording the footage of the area wherein its long way installed. Mostly all of the human identification follows two structures: (1) Face Detection and (2) Face Recognition to acquire the uniformity of person from the live-feed. It includes various statistical and mathematical algorithms for detecting facial features and identifying them through video surveillance. These structures can simply be used to differentiate faces which can be closer to the camera. However, spotting human identity in surveillance camera is not up to the mark to the fact of any of these scenarios like low quality camera, angle of camera, illumination, pose discrepancy, facial expression, unmanageable environment, occlusion, etc. So, Face recognition process is pretty smooth to identify and recognize face from static image but too tough to recognize from video surveillance. Therefore, a manner to recognize spotting human identification in surveillance videos stays very challenging.

Peng et al. [1], presented real-time face tracking algorithm which concentrates on person who has sporting with face mask or without a face mask. Face tracking algorithm includes specially two parts: one is face detector (RetinaFace) and a couple of Kalman filter tracking algorithm (SORT). Retina Face detector involved sub-networks: (1) feature pyramid network (2) independent context module and a couple of multi-assignment module. They use Feature pyramid to take input and produced five-function map to generate multi-task loss. Due to unavailability of video face dataset publicly, they used NRC-IIT facial video dataset and apply face mask on by using different techniques. Analysis result indicates that number of target losses with masks is greater than number of target losses without masks. Proposed system is more accurate on face tracking without a face mask and it lose its tracking while face is protected with face masks.

Rusli et al. [2], Comes up with occlusion problem where person face is covered with some object like face mask in a public place so that face recognition accuracy has been extensively decrease down due to occlusion. Using Multi-Task Cascaded Neural Network (MTCNN) which required equal size of face images and face area has been located. They carried out machine learning LaNet algorithm on dignified training dataset which to fetch the different accuracy of masked and unmasked face. End result which is based on selective strategies indicates that 98.61% recognition accuracy when masked and unmasked each class are trained together. When the masked based dataset images have been taken randomly and tested with trained dataset which achieves average 0.94 accuracy.

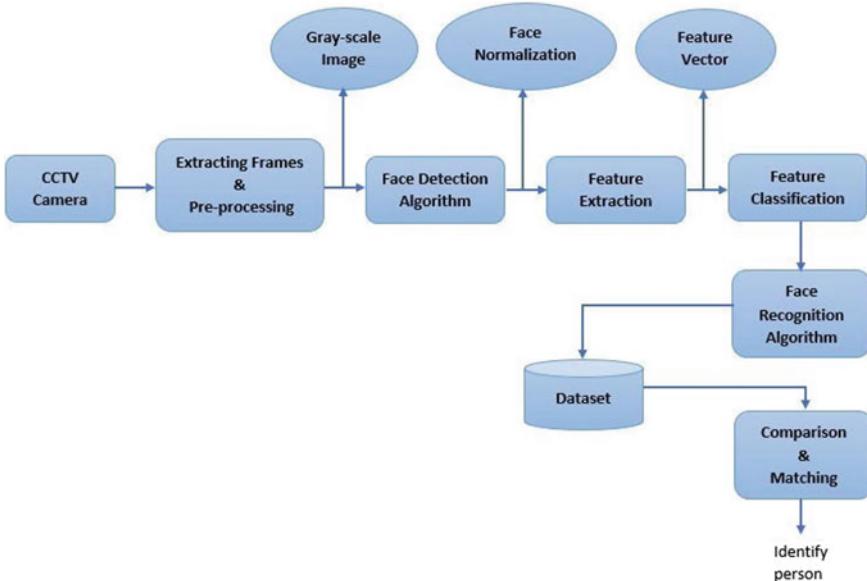
Aswal et al. [3], pursues on Single-camera based face identification which includes mainly two methods: (1) single step and (2) two-step procedure. YOLOv3 model has been used for listing identified people in a single-step and two-step process concerned RetinaFace or YOLOFace that focus on mask faces. VGGFace2 deal with producing facial feature vector for masked based face verification. In addition, face identification process deal with diverse conditions like illumination, orientation, Scale and occlusion. Experimental discussion says that YOLOFace does now no longer work well when faces are cropped, blurred and illuminated with low light and YOLOv3 works precise for cropped, illuminated situation but fails when excessive modifications in person orientation. RetinaFace works excellent in a terms of detection accuracy (0.981), YOLO-Face works proper in a terms of verification accuracy (0.968) and RetinaFace works desirable in terms of overall performance (0.927).

Face reconstruction technique: Malakar et al. [4], observed the problem of face recognition accuracy extensively reduced because of face is protected with a face mask. They used current deep learning methods accompanied by reconstructive techniques to get reconstructed features of occluded feature of the face. Dataset has been built through taking every person image that is in identical length and centered then Principal Component Analysis (PCA) has been carried out to the dataset images to calculate the mean value and mean face image. Then from every decarded image the extracted occluded elements of non-occluded image of a person. Additionally, discards the identical place from the mean and Eigen face image then mean face could be subtracted. Experimental on YALE Face Dataset B indicates that PCA + Deep learning methods offers 85–95% accuracy that is better than conventional Deep learning methods and these methods having a few troubles like all images in a dataset want to be frontal faced and occluded element is not appropriately reconstructed.

Vignesh Kanna et al. [5], comes up with methodology using deep learning techniques via way of means of taking criminal's facial features as well as basic details like gender, shirt or t-shirt pattern and basic apparel as an input for automatic tracking person from the surveillance camera. For training cause Convolutional Neural Network RCNN become used to discover the pattern in an image. After features has been extracted from the image then objects have been identified using Support Vector Machine (SVM) classifier. The confusion matrix has been created to measure the accuracy of this proposed methodology that is 87% and 1.4 average loss of it.

## 2 Proposed System

Primary objective of proposed system is to detect and recognize faces from the surveillance camera to take an automated attendance of students in a classroom as well as detecting people in crowded area like shopping mall, railway station, etc. along with illumination and occlusion. For that Fig. 1 suggests primary steps to carry out face Recognition through CCTV camera. An entire face recognition tool includes four steps as follows: (1) Pre-processing on images (2) Detection of face (3) Extraction



**Fig. 1** Process of face recognition

of Features and (4) Recognizing Face. Firstly, Video frames are extracted from the video and then preprocessing step has been applied for converting RGB image to GrayScale image. This image can also additionally incorporate Background image in addition to face image also. Combine Deep learning techniques with classifier may pertained to discover the face image from the background if present.

Normalization is one of the approaches that will be applied on face image when face has been identified on cycle of face detection. Cropping images, resizing, RGBGray, brightness-adjustment are concerned in a method of normalization to optimize the face popularity. To grow and enhance the overall performance of face popularity in particular operations researcher use: (1) Extraction of features and (2) Classification of features. The feature extraction is an important method to increase the accuracy of face popularity system. Feature vectors are used for illustration of the face in feature space.

### 3 Methodology

Islam et al. [6], proposed a quick and correct system which allows to discover whether the face has been covered with face mask or not. System has 3 stages of implementation: (1) data preprocessing (2) cropping images and (3) image category. System takes input from the Webcam. For the purpose of classification and detection, Convolutional Neural Network (CNN) has been used to detect if face mask is present on

the face or not. Feature extractor and classifiers are the primary key factor of this system in which feature extraction consists of sketched image transformed into new image and this created new image is more precise than sketched image. After that module decreased background noise and additionally implemented a few filtering approaches to reduce the excessive frequency from the image. Result analysis indicates that system generates better accuracy compared with another existing algorithm which is 98%. Main benefit of using this algorithm is system detects the masks face within 0.5 s.

**Identification using Local + global Structure:** Jin et al. [7], proposed a brand-new technique that is used to recognize individual using local as well as global information to obtain the better accuracy in occlusion and deformed condition. Local face information is blended with Global pedestrian detection to acquire the face detection and recognition. Using the MTCNN face detection module, boundary box of detected face has been observed which is used to solve the trouble of detection loss. Feature extraction has been achieved under face recognition using Inception-ResNet-VI network. Experimental result has been carried out at the customized dataset having 93 small videos and frames were extracted from the video and saved in a dataset. Their proposed system accomplished 77.4% of recognition accuracy which is much higher than conventional face recognition system.

**Comprehensive Survey on face detection and recognition:** Sharmila et al. [8] represent the survey on different face recognition algorithms. In which it includes 3 modules including Data Set Creator followed by Trainer and last is Face Recognition. Three algorithms Fisher face, Linear Binary sample Histogram (LBPH) and Eigen Face have been used to detect and recognize the face. Overall achievements of those algorithms have been analyzed using Open-Source Computer Vision Library (OpenCV) in terms of recognition accuracy. According to result multiple face are detected which uses LBPH which achieves 80% of recognition that is much higher than FisherFace and Eigen Face algorithm. But limitation of this look at is that only frontal faces were recognized without any changes in orientation.

All those algorithms have their limitations in which it does not perform better while recognizing face from CCTV camera. Support vector machine (SVM) isn't suitable for massive information Sets. SVM does not performed well when dataset having more noisy data [9]. EigenFaces does not act nicely when faces have the equal face expression and EigenFaces can best come across the front face if there are a few changes in face orientation then this algorithm does not longer carry out higher accuracy [8]. Also, LBPH only detects frontal faces without any variation in face angle [8]. Haar Cascade does now no longer work well in case your face protected with a few objects like glasses, face mask or scarf [10].

After review such research papers, occlusion along with illumination problem required higher attention because recognizing face covered with object in crowded area without illumination is rarely studied. System will used to recognize faces in crowded area which may help in automated attendance in a classroom and identifying people in public place with occlusion and illumination problem. Deep learning techniques helps to achieve higher accuracy in occluded along with illumination

situation. For that firstly from the video frames will be extracted and then using suitable pre-processing methods RGB image will convert into Gray-scale image to make execution faster. Using different dlib's state-of-art built-in libraries along with deep learning, researchers will identify the features of the face and generate face encoding vector. Based on this face encoding vector, researchers will measure the similarities between two images of face and will match that image with dataset images. Using the face recognition algorithm certain measurements on face like size, color, gape between eyebrows and slant of eyes, etc. will be observed. These measurements will be combined to define the face encoding information which will be used to find out the similarities between face images.

### 3.1 Findings

Researchers have identified following findings from above extensive research carried out by different researchers in a field of face recognition from surveillance camera. Combination of pedestrian and face information [7] and convolution neural network [11] can be performed to detect the face. Viola Jones algorithm can be a better choice to detect occluded area of the face [12] and LaNet algorithm can be used to remove undetected part of the face [2]. To achieve higher accuracy of recognition, improved MTCNN [13] algorithm is one of good choice. For feature extraction from the video frames EigenFaces [10] and LBPH feature extraction techniques can be proved as better option to depict the textual of the face images [14].

Co-relation based face detection [15] can be used to measure the response map with CNN to detect the face. MobileNet algorithm [16] can be used to detect the face with face mask which gives better performance compare to KNN and SVM [16]. For face mask detection commercial off-the-shell system [17] and Triplet Loss along with Cosine distance [18] can be better option while detection from video.

Deep learning can be performed best to detect faces from video frames among all techniques. Combined approach can be used with deep learning techniques to increase the recognition rate of the system. However, the method and algorithm to recognize the face from detected frames is highly subjective (Table 1).

## 4 Conclusion and Future Direction

Face detection and recognition has become necessary in everyday routine life because it helps in many areas like security, tracking missing person, criminal identification, automated attendance in universities or colleges. Recognizing face from the surveillance camera is a major problem because of many conditions like illumination, occlusion, camera quality, camera angle, face orientation, face expression and many more. Currently in this covid period, every human face is covered with face mask or video frames taken in public place may suffer with illumination problem or problems like

**Table 1** Analysis on mask based face recognition

No	Theme	Algorithm	Performance	Future scope
1	Analysis performance of machine learning algorithm [8]	Analysis KNN, SVM and MobileNet algorithm	Among these three MobileNet gives excellent result	Can be worked on real data to analysis the actual recognition accuracy
2	Face recognition with face mask [6]	Used MTCNN for occlusion area detection. SVM for face recognition	Recognition rate depends on classifier	Still require improvement in occlusion with illumination problem
3	Face detection using CNN [11]	Feature extraction + CNN for classification and detection	Worked with basically centered image	Will be implemented for the variation in face angle
4	Combination of PCA + Deep learning [4]	PCA algorithm applied after discarded occluded part	Only worked with frontal face dataset	Will work on different face angle and occlusion
5	Evaluation of mask and unmask face [2]	Used MTCNN for finding regions and LaNet algorithm for recognition	Best for unmasked faces compared to mask face	Will apply machine learning to improve result
6	Real-time face tracking with face mask [1]	Used Kalman filter for face detection	Failed if majority of face covered with some object	Will be used machine learning on some real face dataset
7	Cosine distance based face detection [15]	Used Haar-cascade for face detection. Triplet loss and VGG are used for recognition	VGG is good option com-pare to triplet loss in terms of execution time	Used this method with illumination
8	Single camera mask detection [3]	Used YOLOv3 and RetinaFace model for detection. VGGFace2 for recognition	VGGFace2 and RetinaFace works well with occlusion	Will replace VGG-Face2 with ArcFace to obtain more accuracy

overlapping images. This research review is further useful for researchers to develop automatic application to recognize faces from surveillance camera for public place to solve occlusion along with illumination problem in public place.

## References

1. X. Peng, H. Zhuang, G.-B. Huang, H. Li, Z. Lin, Robust real-time face tracking for people wearing face masks, in *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)* (2020), pp. 779–783. <http://doi.org/10.1109/ICARCV50220.2020.9305356>

2. M.H. Rusli, N.N.A. Sjarif, S.S. Yuhaniz, S. Kok, M.S. Kadir, Evaluating the masked and unmasked face with LeNet algorithm, in *2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA)* (2021), pp. 171–176. <http://doi.org/10.1109/CSPA52141.2021.9377283>
3. V. Aswal, O. Tupe, S. Shaikh, N.N. Charniya, Single camera masked face identification, in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)* (2020), pp. 5760. <http://doi.org/10.1109/ICMLA51294.2020.00018>
4. S. Malakar, W. Chiracharit, K. Chamnongthai, T. Charoenpong, Masked face recognition using principal component analysis and deep learning, in *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (2021), pp. 785–788. <http://doi.org/10.1109/ECTI-CON51831.2021.9454857>
5. J.S. Vignesh Kanna, S.M. Ebenezer Raj, M. Meena, S. Meghana, S. Mansoor Roomi, Deep learning based video analytics for person tracking, in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)* (2020), pp. 1–6. <http://doi.org/10.1109/icETITE47903.2020.173>
6. M.S. Islam, E. Haque Moon, M.A. Shaikat, M. Jahangir Alam, A novel approach to detect face mask using CNN, in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)* (2020), pp. 800–806. <http://doi.org/10.1109/ICISS49785.2020.9315927>
7. K. Jin, X. Xie, F. Wang, X. Han, G. Shi, Human identification recognition in surveillance videos, in *2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (2019), pp. 162–167. <http://doi.org/10.1109/ICMEW.2019.00-93>
8. Sharmila, R. Sharma, D. Kumar, V. Puranik, K. Gautham, Performance analysis of human face recognition techniques, in *2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)* (2019), pp. 1–4. <http://doi.org/10.1109/IoT-SIU.2019.8777610>
9. S. Saypadith, S. Aramvith, Real-time multiple face recognition using deep learning on embedded GPU system, in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* (2018), pp. 1318–1324. <http://doi.org/10.23919/APSIPA.2018.8659751>
10. T. Mantoro, M.A. Ayu, Suhendi, Multi-faces recognition process using Haar cascades and Eigenface methods, in *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)* (2018), pp. 1–5. <http://doi.org/10.1109/ICMCS.2018.8525935>
11. S. Kakarla, P. Gangula, M.S. Rahul, C.S.C. Singh, T.H. Sarma, Smart attendance management system based on face recognition using CNN, in *2020 IEEE-HYDCON* (2020), pp. 1–5. <http://doi.org/10.1109/HYDCON48903.2020.9242847>
12. A. Raghunandan, Mohana, P. Raghav, H.V.R. Aradhya, Object detection algorithms for video surveillance applications, in *2018 International Conference on Communication and Signal Processing (ICCSP)* (2018), pp. 0563–0568. <http://doi.org/10.1109/ICCSP.2018.8524461>
13. C. Qi, L. Yang, Face recognition in the scene of wearing a mask, in *2020 International Conference on Advance in Ambient Computing and Intelligence (ICAACI)* (2020), pp. 77–80. <http://doi.org/10.1109/ICAACI50733.2020.00021>
14. N. Gupta, P. Sharma, V. Deep, V.K. Shukla, Automated attendance system using OpenCV, in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (2020), pp. 1226–1230. <http://doi.org/10.1109/ICRITO48877.2020.9197936>
15. H.W. Hsu, T. Wu, W.H. Wong, C.Y. Lee, Correlation-based face detection for recognizing faces in videos, in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2018), pp. 3101–3105. <http://doi.org/10.1109/ICASSP.2018.8461485>
16. W. Vijitkunsawat, P. Chantngarm, Study of the performance of machine learning algorithms for face mask detection, in *2020—5th International Conference on Information Technology (InCIT)* (2020), pp. 39–43. <http://doi.org/10.1109/InCIT50588.2020.9310963>

17. N. Damer, J.H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, A. Kuijper, The effect of wearing a mask on face recognition performance: an exploratory study, in *2020 International Conference of the Biometrics Special Interest Group (BIOSIG)* (2020), pp. 1–6
18. D.A. Maharani, C. Machbub, P.H. Rusmin, L. Yulianti, Improving the capability of real-time face masked recognition using cosine distance, in *2020 6th International Conference on Interactive Digital Media (ICIDM)* (2020), pp. 1–6. <http://doi.org/10.1109/ICIDM51048.2020.9339677>

# An Innovative Framework Based Algorithmic Approach for Object Detection



Anshika Jain and Maya Ingle

**Abstract** Object Detection techniques have been enumerable used in a variety of real-time applications such as robotics, crime investigation, transportation, etc. Object classification is an important task in computer vision and is the process of tagging objects into predefined and semantically significant classes using trained datasets. A framework for object detection using the Viola-Jone algorithm for object classification is proposed in this paper. The architecture of the framework encompasses image acquisition, image pre-processing, classification and extraction of objects, and computation of related measures. The experiment is performed on a dataset containing 120 images of human faces with different angles, poses, and light conditions. It is worth mentioning that the faces as objects are recognized successfully from the set of input images with a rate of 96.67%. Moreover, the objects on faces such as eyes, nose, and mouth are detected successfully with an average accuracy of 93.10%, 93.10%, and 90.80% respectively. The attributes/measures of these objects are vital for computer recognition and hence, the properties/measures of respective objects are computed at the end. This framework will be useful in several real-life applications; especially in criminal investigation applications for identification of persons/criminals and computer portrait designing, etc.

**Keywords** Object detection · Image processing · Viola-Jones algorithm

## 1 Introduction

With the increasing need for identity-based security and industrial applications, object detection has become attractive among the researcher community. Object detection algorithms use extracted features and learning algorithms to identify instances of object categories. Continuous efforts are visible on various techniques of object detection along with their implementation in real-time applications. Some object detection methods focus upon the detection of the frontal face with sufficient

---

A. Jain (✉) · M. Ingle  
SCS&IT, DAVV, Indore, India  
e-mail: [anshu261293@gmail.com](mailto:anshu261293@gmail.com)

light conditioning and camera quality. The main motive is to build a system responsible to detect and recognize the textures of human parts of the body in an image or a video. The estimated parameters of the parts in the human body are tracked with the various parameters of facial features. Face Detection is a challenging task as it deals with the identification of different sizes, shapes, textures, and varying intensities of colors in it [1]. This can be further applied to real-world applications such as online exams, identifying the persons' gender/age, etc. The applications of object detection cover mainly the domains of Computer Science such as Human–Computer Interaction (HCI), Robotics, Consumer Electronics, Security, Transportation, etc. [2]. Some complex object detection approaches such as statistics, fuzzy logic, neural networks, etc. have proven to be useful. A robust and 15 times quicker than any technique exists reporting 95% accuracy for face detection at the time of release [3]. Object classification is an important task in computer vision and is the process of tagging objects into predefined and semantically significant classes using trained datasets. Tensor features with Scale Invariant Feature Transform (SIFT) have been used to improve the accuracy of the task [4]. However, an innovative framework may be developed for object detection with a higher success rate using an algorithmic approach for object classification.

## 2 Proposed Framework

In this paper, an innovative framework for object detection based on the Viola-Jone Algorithm is presented. It specifically works on the extraction of the face along with existing objects on it such as eyes, nose, and mouth mainly. The architecture of the framework encompasses image acquisition, image pre-processing, classification and extraction of objects, and computation of related measures. Image Acquisition is concerned with capturing images either with browsing images or capturing the real-time images using the camera. The input images may exist in a variety of formats such as .jpg, .png, .tif, .jpeg etc. Image pre-processing deals with grayscale conversion, noise removal, and filtering. Grayscale Conversion is responsible for enhancing the feature of the input image by reducing unwanted distortions. During grayscale conversion, an input RGB image is transformed into gray-level color space. The grayscale conversion usually reduces the strength of pixel intensity and represents an image into black and white color with numerous darks of gray existing without them. An input RGB image comprises three-dimensional complexities. Gray level conversion transforms three-dimensional complexities into one-dimensional form to reduce the asset of each pixel and represents the minute features of an image appropriately. An input RGB image includes an 8-bit pixel value for each color channel in the range of 0–255. During color space conversion, the gray level image represents a single 8-bit pixel strength for individual pixels in an image [5]. Noises occur during image acquisition or digitization or image transmission and are categorized into various forms such as Impulse noise (Salt-and-Pepper noise), Amplifier

noise (Gaussian noise), Additive noise, Quantization noise (Uniform noise), Multiplicative noise (Speckle noise), and Poisson Noise (Photon noise) in general [6, 7]. In image processing, filtering is a basic function used to perform noise reduction, smoothing, polishing, and re-sampling to obtain the true pixel value from the noisy image. Filtering techniques are mainly classified as Linear filters includes Mean filter, Gaussian filter, and Wiener filter, whereas Non-Linear filters cover Median filter, Corner filter, and LoG filter. Literature reveals that the Median filter performs the best with Gaussian noise on average [8]. A framework based on the Viola-Jones Object Detection algorithm is proposed focusing upon the classification of objects in real-time with higher accuracy and the computation of respective measures as shown in Fig. 1. It comprises Integral Image, Adaboost, and Cascade Classifier. Integral

**Fig. 1** Proposed framework for object detection

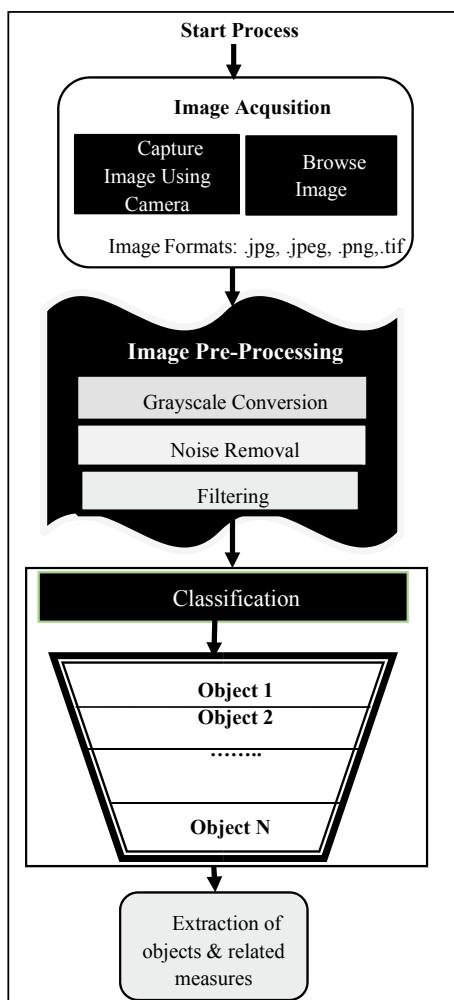


Image is concerned with the identification of objects using sub-window whereas Adaboost is used to check the suitability/correctness of identified image in the rectangle. Cascade Classifier works sequentially until the best suitable match of the input image is found [9–11]. The detected objects by Cascade Classifier are stored in a specific database. The attributes of these objects are vital for computer recognition and hence, the properties/measures of respective objects are computed.

### 3 Algorithmic Approach of Framework

In this section, an algorithmic approach of our framework is presented.

**//Algorithm: Computation of measures for n input images using a proposed framework//**

```

Step I: Start { begin i= 1 to n
Step II: //Acquisition and Conversion of an Input Image //
    II(i)=read_image_from_jpg
    GI(i) =rgb2gray(II(i))
Step III: //Salt & Pepper Noise removal and Median Filtering //
    NI(i)= apply_noise(GI(i))
    FI(i) =apply_filter(NI(i))
Step IV: //Identification of Objects //
    Detect=Cascade_object_from_filtered image(FI(i))
Step V: //Identification and cropping of object of interest in rectangle //
    Cropped_Image=Idfy_crop_object(Detect)
Step VI: // Saving extracted object in database //
    Save_image (Cropped_Image in Output_file)
Step VII: //Computation of m Measures //
    M1=Save_image_edge_row(max-min)
    M2= Save_image_edge_coloum(max-min)
    M3, M4, ..... Mm
Step VIII: end } End

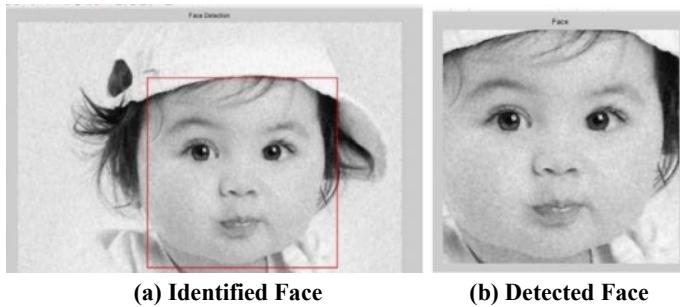
```

### 4 Experimental Execution

For the execution of our framework, we use Matlab 2013a with 4 GB RAM, 64 Bit Windows Operation System. The experiment is performed on a dataset containing  $n = 120$  images of human faces with different angles, poses, and light conditions. These images are made available for experimentation from Labeled Faces in the Wild (lfw), a public benchmark for face verification. These images are basically colored images of size i.e.  $2 \times 2$  and acquired in jpg format for further processing.



**Fig. 2** Images at various stages



**Fig. 3** Output images

Using suitable functions of Matlab 2013a, the conversion of these rgb images into grayscale is performed. The noise existing in grayscale images has been eliminated by Salt & Pepper noise removal and Median filtering techniques. Figure 2a–c depicts the colored input image, its converted grayscale image, and the filtered image respectively. Now, it is possible to identify the objects of interest by applying the Cascade classifier. In our experimentation face, eyes, nose, and mouth are considered as objects. Figure 3a and b represents the identified object as a face along with its cropped version respectively. These cropped images are stored in the output file for further computations. Lastly, the number of properties/measures of interest (in our case  $m = 2$ ) such as height and width of objects in our case are computed.

## 5 Results and Conclusion

For experimentation, we have used 120 input images and 4 objects as specified earlier. After execution of our algorithmic approach in the framework, it is worth mentioning that the faces as objects are recognized successfully from the set of input images with a rate of 96.67%. Moreover, the objects on faces such as eyes, nose, and mouth are detected successfully with an average accuracy of 93.10%, 93.10%, and 90.80% respectively as shown in Table 1. Their average measures (Height and Width) have been computed in terms of pixel values are also shown in Table 1. An innovative

**Table 1** Experimental results of object detection

Objects in images	Outcome of <i>Viola-Jones</i> algorithm		Computation of measures	
	No. of images (total = 120)	Success rate of recognition (%)	Average height (M <sub>1</sub> ) (pixel val)	Average width (M <sub>2</sub> ) (pixel val)
Face	116	96.67	220.68	211.02
Eyes	112	93.10	28.51	61.91
Nose	112	93.10	69.40	56.30
Mouth	109	90.80	37.95	75.40

approach has been used along with the Viola-Jones algorithm for object identification and computation of measures with a good success rate. Thus, we conclude that this framework may be useful in several real-life applications; especially in criminal investigation applications for the identification of persons/criminals and computer portrait designing, etc.

## References

1. M.K. Dabhi, B.K. Pancholi, Face detection system based on Viola-Jones algorithm. *Int. J. Sci. Res. (IJSR)* **5**(4) (2016). ISSN (Online): 2319-7064
2. Q.-r. Jiang, H.-l. Li, Robust human face detection in complicated color images, in *2nd IEEE International Conference on Information Management and Engineering (ICIME)* (IEEE, 2010)
3. M.U. Riaz, D.S. Benini, Face Anonymiser for Images and Videos. Anno Accademico 20012/2013
4. N. Najva, SIFT and tensor based object classification in images using Deep Neural Networks, in *Proceedings of International Conference on Information Science (ICIS)*, Kochi, India (2016), pp. 32–37
5. T. Kumar, K. Verma, A theory based on conversion of RGB image to gray image. *Int. J. Comput. Appl.* (0975-8887) **7**(2) (2010)
6. S. Kaur, Noise types and various removal techniques. *Int. J. Adv. Res. Electron. Commun. Eng.* **4**(2) (2015)
7. M.D. Sontakke, M.S. Kulkarni, Different types of noises in images and noise removing technique. *Int. J. Adv. Technol. Eng. Sci.* **3**(1) (2015)
8. A. Malhotra, E.D. Kumar, Image denoising with various filtering techniques. *Int. J. Adv. Res. Comput. Sci. Technol.* **3**(2) (2015)
9. C. Lin, Face detection by color and multilayer feedforward neural network, in *IEEE International Conference on Information Acquisition*, Hong Kong and Macau, China (2005), pp. 518–523
10. K. Vikram, S. Padmavathi, Facial parts detection using Viola Jones algorithm, in *International Conference on Advanced Computing and Communication Systems (ICACCS-2015)*, IEEE, Coimbatore, India, 06–07 Jan 2017, 978-1-5090-4559-4/17/\$31.00
11. E. Alionte, C. Lazar, A practical implementation of face detection by using Matlab cascade object detector, in *19th International Conference on System Theory, Control and Computing (ICSTCC)*, 14–16 Oct 2015

# Smart Bus Concept of COVID-19 Prevention Using FaceBio Keys



Dhara Rathore, Akash Saxena, and Gaurav Kumar Das

**Abstract** The COVID-19 surge has impacted surge has an impact of each and every part of world. Vaccination is going on but still some measures are required to maintain the social distances to reduce transmission of virus. The transportation is the rapid source of transmission of diseases as daily number of people makes use of public transportation. We have proposed the concept of the smart bus will can contribute a lot in controlling its massive spread. So, we will be proposing the new design for the smart bus for promoting the contact-less journey. In the case, the door entry will be based on the face recognition concept, for the purpose of the face recognition, we will use of the concept of the facial landmark analysis and with the concept of the machine learning. For the transportation purpose, the face parameters are linked with the smart bus database, which require the passenger to take a pre-paid card, in order to enjoy the travel. The bus will have the camera which recognize in the passenger face and give the entry and the bus will be automated driving mode using the google navigation system and traffic sensing system. The bus will be driver less and the bus will have sensors in the seats which operate on the pressure concept, so the seating arrangement in the bus is also on basis of the social distancing. If the social distancing, is not maintained as passenger tries to sit on the wrong seat and the alarm them will sound in the bus and the bus will not move further. So, using these concepts, we can reduce the COVID-19 chances of spreading due to the usage of public transportation services.

**Keywords** Smart bus · COVID-19 · Social distancing · Face recognition

## 1 Introduction

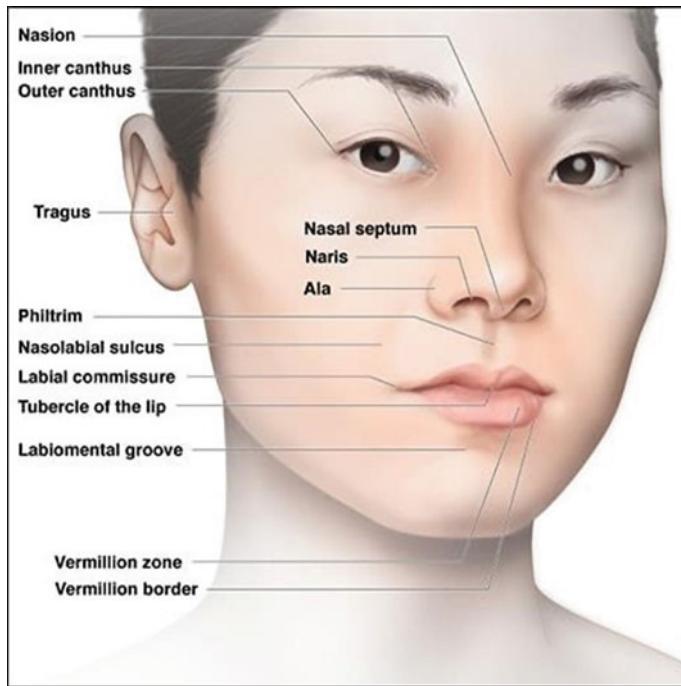
Face Arrangement settle for a vital activity in face acknowledgment, outward look examination, facial action, 3D face amusement and facial icon change of state. Facial

---

D. Rathore (✉) · A. Saxena · G. K. Das

Department of Computer Science and Engineering, Compucom Institute of Technology and Management, Jaipur, India

e-mail: [dhararathore1@gmail.com](mailto:dhararathore1@gmail.com)



**Fig. 1** Facial landmark

land-marks, usually known as facial half centers or fiducially facial centers area unit organized at linguistics zones, for example, eye structures, eye brows, mouth shapes, space of facial landmarks could be a set of face game set up frameworks. This survey regarding signally about landmark speech act and following mistreatment vital learning (Fig. 1).

## 2 Literature Survey

**Song et al.** [1] In this paper, a profound engineering for thick facial landmark identification, called ConcatNet, is then proposed. In design that the creators proposed, CNN-put together thick landmark locator with respect to part districts of the face, which expands a given arrangement of meager landmarks to increasingly precise and thick landmarks.

**Baddar et al.** [2] In this paper, creators propose another facial landmarks location technique dependent on profound learning with facial shape and facial segments imperatives. The proposed profound convolutional neural systems (DCNNs) for facial landmark discovery comprises of two profound systems: one DCNN is to

recognize landmarks obliged on the facial form and the other is to identify landmarks compelled on facial segments.

**Wei et al.** [3] Facial demeanor examination assumes a significant job in nonverbal vivid interchanges. Considering facial landmark confinement is a fundamental part for outward appearance investigation, it draws a lot of consideration. Since the precision of facial landmark confinement is connected with various appearances and is influenced by the consequence of face discovery, creators propose a two organize facial landmark limitation strategy in this paper. Trial results on the broaden Cohn-Kanade database show that the proposed technique has increasingly exact facial landmarks confinement result.

**Lee et al.** [4] Facial landmark identification is then significant earlier data for the other face arrangement related issues, for example, head estimation, facial feeling appearance, and face demonstrating. Among the ordinary facial landmark discovery algorithms, the TREE algorithm which utilizes the fell relapse technique can recognize facial landmark much quicker than the different algorithms utilizing little preparing information.

**Choi et al.** [5] Emotion acknowledgment dependent on outward appearances is significant for cooperation among human and computerized reasoning (AI) framework, for example, social robots. In this paper, creators propose a two-dimensional (2D) landmark highlight for viably perceiving facial smaller scale appearance..

**Guo and Qi** [6] Facial element parsing is a functioning exploration point in picture understanding. In this paper, creators address the element parsing issue by applying low-rank grid disintegration on facial pictures.

**Terada et al.** [7] Various applications have been created in facial examination to help the acknowledgment of individual ascribes, investigation of race, and individual validation for the security business and other exploration fields. Accordingly, it is feasible to recognize the distinctions fit as a fiddle dependent on spot of birth or country. The proposed concept makes use of the 3-D pictures to perform the facial based analysis and also then made the use of the CNN Based approach for the analysis purpose.

**Özseven and Düğenci** [8] also, the characterization was then performed with the assistance of these components and the most elevated achievement was then found with the 94.60% with the use of the MLP classifier. Acquired discoveries show the ease of use of the distance and slant between the milestones in facial acknowledgment.

**Happy and Routray** [9] Extraction of discriminative highlights from striking facial patches assumes a fundamental job in compelling outward appearance acknowledgment. The exact discovery of facial landmarks one that improves the restriction of the notable fixes on face pictures. This paper thus proposes the novel structure for demeanor acknowledgment by utilizing appearance highlights of chosen facial patches.

**Glory et al.** [10] introduced the novel approach of creating passwords using random text extraction. The information provided or obtained from the user forms the foundation for the password, such as the user's favorite novel name, the number of grad mother's children, key dates, and so on. The text is retrieved at random from the specified information as well as strong passwords.

### 3 Proposed Work

The proposed work involves the following modules.

#### 3.1 New Registration for Bus

This section will discuss the new registrations for the bus passengers, the registration is based on the identification of user using Face Recognition, Finger Print and smart pattern.

- Step 1: Read Passenger Name, Phone or Email Id and Address
- Step 2: If Phone/Email ID already in Use Then:
  - i. Write "Passengers Details Exists in Database"
  - ii. Goto End
- [End of If Structure]
- Step 3: Input Face Image, generate the SHA-512 Code and store in FACESHA.
- Step 4: Input BIO-Metric Finger Print Image, generate the SHA-512 Code and store in BIOSHA.
- Step 5: Select Smart Pattern clicking on Button and click on generate Pattern.
- Step 6: Extract 20 characters FACESHA and concatenate Smart Pattern in order to generate FACEKEY
- Step 7: Extract 20 characters BIOSHA and concatenate Smart Pattern in order to generate BIOKEY
- Step 8: Stores the details in BusRegis table
- Step 9: END.

#### 3.2 Bus Entry Process

This section is concerned about the entry in the bus by the passenger using Face or Fingerprint validation

Step 1: Read Passenger Name , Phone or Email Id

Step 2: If Validation using Face then

- a. Input Face Image, generate the SHA-512 Code and store in FACESHA.
- b. Select Smart Pattern clicking on Button and click on generate Pattern.
- c. Extract 20 characters FACESHA and concatenate Smart Pattern in order to generate AUTHKEY

Else

- a. Input BIO-Metric Finger Print Image , generate the SHA-512 Code and store in BIOSHA.
- b. Select Smart Pattern clicking on Button and click on generate Pattern.
- c. Extract 20 characters BIOSHA and concatenate Smart Pattern in order to generate AUTHKEY

[End of If Structure]

Step 3: If valid Phone or Email Id and AUTHKEY then

Allow in Bus

Else

Denied from Bus Entry

[End of If structure] Step 4:

END.

### **3.3 Seat Occupancy Module**

Step 1: In the case of the seat module we can use the sensors on the seats of the bus in order to detect the occupancy of the seat, and according to the social distancing plan if the invalid seat is occupied the alarm will beep up.

Step 2: The Seat Occupancy sensor one which is then magnetically operated by making use of the push button sensor with the simple push-fit clip mounting configuration, that will allows the passenger safety system then to determine or detect the presence of an object or the person occupying the seat.

### **3.4 Bus Floor Module**

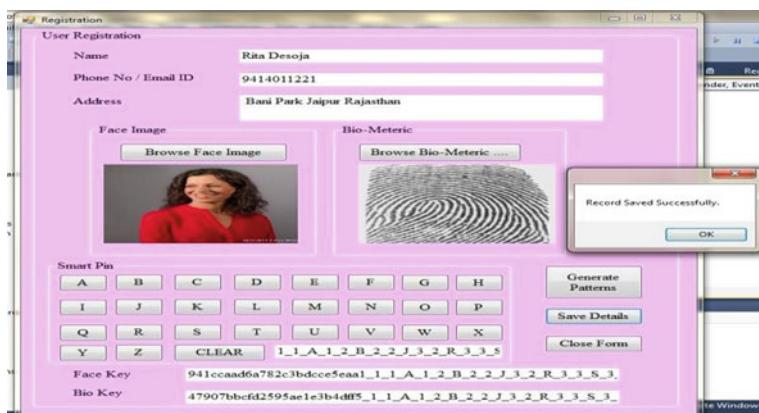
Now when the driver has to start the bus the smart sensors on the bus floor movement area can detect the presence of any standing passengers so that module can stop the bus from starting so the driver can identify the presence of the extra passenger so standing passengers can be restricted in this way.

## 4 Implementation

In the Registration section we have in the entry module we considering about the entry of passengers in the bus, so here we have suggested the entry based on the face recognition and finger print recognition. The process starts with the entry of the passengers details like Name, phone or Email Id and Address of Passenger. Then the user has to Input Face Image, then generated the SHA-512 Code and that is stored in FACESHA. After that user has to Input BIO-Metric Finger Print Image, then generated the SHA-512 Code and stored in BIOSHA.

After that select Smart Pattern clicking on Button showing Alphabets and click on generate Pattern. For the formation of the FACEKEY first Extract 20 characters from the FACESHA and concatenate Smart Pattern in order to generate FACEKEY. For the formation of the BIOKEY Extract 20 characters BIOSHA and concatenate Smart Pattern in order to generate BIOKEY.

Figure 2 shows the Passenger Registration form, after the whole process the details are stored in the BusRegis database table which is used for storing the details of the registered passengers. The structure of the BusRegis table stored the following information name of passenger, Phone or Email ID, Address of Passenger, Path of Face Image, Path of Finger Print Image, SHA-512 of Face Image, SHA-512 of Finger Print, Smart Code Pattern, the face key and also the biokey. In the case of the seat module we can use the sensors on the seats of the bus in order to detect the occupancy of the seat, and according to the social distancing plan if the invalid seat is occupied the alarm will beep up. The Seat Occupancy sensor one which is then magnetically operated by making use of the push button sensor with the simple pushfit clip mounting configuration, that will allows the passenger safety system then to determine or detect the presence of an object or the person occupying the seat (Fig. 3).



**Fig. 2** Registration process



**Fig. 3** Seat module

## 5 Result and Discussion

**Base sample generated Pattern is [10] @niK29tUr@n?**

And the proposed concept form the pattern FaceKey and BIOKey,

**FaceKey:** ba0d408aeeb0213de25179ff9\_1\_1\_A\_2\_2\_J\_3\_3\_S\_2\_4\_L\_1\_5\_E\_

**BIOKey:** 88a1d01e1e4c0a3403ce186e4\_1\_1\_A\_2\_2\_J\_3\_3\_S\_2\_4\_L\_1\_5\_E\_

In order to check the strengths of the passwords of the base and the proposed approach we have used various types of tools to check the entropy or the time required to break the password pattern.

Thus, we have divided the concept as

- Time Required to Break the Password.
- Entropy—Strength of the password.

### 5.1 On Basis of Time Required

#### 5.1.1 My1Login

This tool calculates the strength of the password on the basis of the years required to break the password (Tables 1 and 2).

**Table 1** Strength checking tool 1-years required comparison with FaceKey

	Base approach	Proposed FaceKey
Years taken to break	63 million years	192 trillion trillion trillion trillion years

**Table 2** Strength checking tool 1-years required comparison with BioKey

	Base approach	Proposed FaceKery
Years taken to break	63 million years	8 trillion trillion trillion trillion trillion trillion years

**Table 3** Strength checking tool 2-years required comparison with FaceKey

	Base approach	Proposed FaceKey
Years taken to break	4 hundred thousand years	38 billion quinquagintacentillion years

**Table 4** Strength checking tool 2-years required comparison with BioKey

	Base approach	Proposed FaceKery
Years taken to break	4 hundred thousand years	3 octovigintillion years

### 5.1.2 Security.org

This tool calculates the strength of the password on the basis of the years required to break the password (Tables 3 and 4).

## 6 Conclusion

COVID-19 is the pandemic which arises in 2020 and effected all the countries around the world. The main issue with the Corona virus is the community spread and public mode of transportation is main source of spread as seen in the first as well as more important in the second phase of impact of corona virus. The proposed concept work on proper authentication of Bus Entry on the basis of FaceKey and BIOKey which are generated using the proposed algorithm and compared with the previous approaches for the strength via online tools and results are more effective in proposed work.

## References

1. H. Song, J. Kang, S. Lee, ConcatNet: a deep architecture of concatenation-assisted network for dense facial landmark alignment, in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Oct 2018, pp. 2371–2375
2. W.J. Baddar, J. Son, D.H. Kim, S.T. Kim, Y.M. Ro, A deep facial landmarks detection with facial contour and facial components constraint, in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, Sept 2016, pp. 3209–3213
3. W. Wei, C. Tian, Y. Zhang, A two-stage facial landmark localization method, in *2015 International Conference on Orange Technologies (ICOT)*, Hong Kong, Dec 2015, pp. 157–160

4. Y. Lee, T. Kim, T. Jeon, H. Bae, S. Lee, Facial landmark detection using Gaussian guided regression network, in *2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CS&C)*, JeJu, Korea (South), June 2019, pp. 1–4
5. D.Y. Choi, D.H. Kim, B.C. Song, Recognizing fine facial micro-expressions using two-dimensional landmark feature, in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Oct 2018, pp. 1962–1966
6. R. Guo, H. Qi, Facial feature parsing and landmark detection via low-rank matrix decomposition, in *2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Sept 2015, pp. 3773–3777
7. T. Terada, Y. Chen, R. Kimura, 3D facial landmark detection using deep convolutional neural networks, in *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Huangshan, China, July 2018, pp. 390–393
8. T. Özseven, M. Düğenci, Face recognition by distance and slope between facial landmarks, in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Sept 2017, pp. 1–4
9. S.L. Happy, A. Routray, Automatic facial expression recognition using features of salient facial patches. *IEEE Trans. Affect. Comput.* **6**(1), 1–12 (2015)
10. F.Z. Glory, A. Ul Aftab, O. Tremblay-Savard, N. Mohammed, Strong password generation based on user inputs, in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019, pp. 0416–0423

# Data Integration Solutions—A Key to the Cloud Migration



Ritul Bakshi and Vijay Singh Rathore

**Abstract** The health industry produces a large amount of data on daily basis. Within this vast variety of data lies the valuable insights, which, when applied judiciously can bring in wonderful results. Managing these ‘Healthcare Big data’ using traditional tools is near to impossible. These challenges can be resolved by migrating the data to the cloud. Cloud migration has given rise to various tools to deal with different cloud migration use cases. However, because multiple platforms and tools have been provided by different cloud computing providers, so it is often difficult for the Data Engineers to understand and select the right tools that address their cloud migration use case. This paper aims towards looking for the right solution for migrating a health data analytical tool from the database to the cloud by performing a comprehensive comparison between the available solution and ensuring that the solution should be good enough to cater to analytical enhancements.

**Keywords** Cloud computing · Cloud providers · Data integration solutions

## 1 Introduction

Technology always moves towards connectivity. To connect systems, databases, and applications to the way we need them to, Data Integration Solutions are very helpful. Any application whether it is a legacy system or database whose contents are siloed off or an enterprise with a hundred disparate SaaS applications, SOA, and APIs, Data integration Solution solves the modern connectivity problem in a scalable way.

Cloud data integration can be referred to as a practice that integrates data that is further used by different systems. The integration is performed within public or private clouds, or it can be performed between cloud-based and on-premises systems. Unified data stores can be created using a data integration solution.

---

R. Bakshi (✉) · V. S. Rathore  
IIS (Deemed to be University), Jaipur, Rajasthan, India  
e-mail: [ritul.bakshi@gmail.com](mailto:ritul.bakshi@gmail.com)

## 2 Data Integration Solution

### 2.1 *On-premise Data Integration Solutions*

These solutions are used to integrate data from different on-premises data sources. In order to use these solutions, it needs to be installed in a private cloud or over the local networks with optimized native connectors. The connectors will help in enabling the batch loading from different data sources.

### 2.2 *Cloud Based Data Integration Solutions*

These solutions are used to integrate data and then use that further in the cloud. These solutions perform data integration using different data sources and can be used as an integration platform as a service.

### 2.3 *Open-Source Data Integration Solutions*

These solutions are used to replace any proprietary and potentially expensive enterprise solutions. These solutions provide complete control of the data.

### 2.4 *Proprietary Data Integration Solutions*

These tools have different pricing as compared to open-source tools. They are generally built for specific business use cases.

## 3 Different Cloud Based Data Integration Solutions

### 3.1 *Microsoft Azure Data Factory*

Azure Data Factory is a bit more than only an ETL tool. Data Factory allows users to create data-driven workflows for orchestrating data movement and transformation at scale [1]. Westin and Stendal [2] performed a case study of the implementation of Microsoft Azure in the power industry.

### **3.2 AWS (*Amazon Web Services*) Glue**

AWS (Amazon Web Services) Glue has visual as well as code-based interfaces. AWS (Amazon Web Services) Glue Studio is used by Data engineers and ETL (extract, transform, and load) developers to create, run, and monitor ETL workflows. AWS (Amazon Web Services) Glue is a serverless ETL that allows users to automate data preparation and analytics [1].

### **3.3 IBM (*International Business Machines*) DataStage**

IBM (International Business Machines) DataStage uses a container-based architecture. Large workloads can be handled by deploying IBM DataStage and Information Server. The balanced optimization available in DataStage offers performance improvements for relational data sources, however, its characteristics on NoSQL and stream DSS are unknown [3].

### **3.4 *Informatica PowerCenter***

As per Gupta and Sahayadhas [4] to perform ETL or ELT operations, there are so many tools are available in the market and out of which Informatica is the ETL tool preferred by most of the IT companies, through which these operations can be done easily.

### **3.5 *Oracle Data Integrator***

Oracle Data Integrator has interoperability with Oracle Warehouse Builder (OWB) that's why migration for OWB customers is very easy. As per Wycislik et al. [5], ODI (Oracle Data Integrator) simplifies and improves the efficiency of most common ETL tasks.

## **4 Review of Cloud Based Data Integration Solutions**

A recent report by Seagate and IDC found that while enterprise data is projected to increase at 42.2% annually over the next two years, up to 68% of it may go unused [6].

#### **4.1 Literature Reviews**

- Biplob et al. [7] performed a comparison of different existing tools and tried to suggest the best solution for different scenarios.
- Mukherjee et al. [8] compared different properties of some popular tools and highlighted their advantages and disadvantages.
- Diou et al. [9] reviewed different solutions by highlighting the performance aspects of data processing in terms of volume and velocity.
- Jayashree et al. [10] initiated a discussion on different data integration methods and discussed the changes faced by them while processing the XML data using the Extraction, Transformation, and Load method.
- Sreemathy et al. [11] have explained the integration using the Talend Open Studio ETL tool and performed an analysis of the ETL method.
- Khan et al. [12] have discuss an ETL process that allows users to remodel and store data, and a proposed schema for a Common Data Model which will house this transformed data under one roof [5].
- Sabtu et al. [13] discussed the challenges of implementing ETL system for streaming or near real-time data which needs to evolve and streamline itself with the different requirements and also presented the current efforts and solution approaches to address the challenges [12].

#### **4.2 Findings**

- Data Integration tools are very important for business intelligence and using a proper ETL tool we can get the desired outcome easily.
- By doing the proper preparation cloud computing barriers can be easily overcome.
- Serverless moves us towards continuous development and delivery.
- The common use cases for each data integration solution have not been discussed in any of the research papers.
- New integration solution that is based on a dedicated programming environment gives better performance as compared to graphical tools.
- Traditional ETL needs to be enhanced to support streaming of data processes with federation across multiple sources.
- Any solution has not been provided for the business use case of migrating a legacy database application to the cloud with the help of a data integration solution.

#### **4.3 Limitations of the Existing Research**

- Existing research did not highlight the most common use cases for each data integration solution.

- Existing research does not present a thorough analysis from the perspective of Health Data Analytics.
- Existing research does not discuss the scalability and documentation support of the available solutions.
- Existing research does not cover that which solution provides a better personalized customer experience.
- Existing research does not discuss in detail about different types of data integration solutions.
- Existing research does not give details about the usability of the parameters that have been selected for the comparison.

## 5 Comparative Analysis of Different Cloud Based Data Integration Solutions

Many data integration tools are available, so it is very difficult to compare all the tools that are available in the market. Based on the study of different blogs, surveys five most efficient solutions have been compared. Table 1 compares different aspects of the comparison among different solutions.

## 6 Result and Discussion

The review has been performed for 5 different data integration solutions. This paper contains informative information that will help readers while performing each step of cloud migration. For example, identifying services used in the existing application and the services on which the existing application depends is the first step for any migration and this paper contains the details of the connectors supported by each integration solution so the reader can check that which solution support the required services. Similarly, for migration, the developer needs to decide that whether they are going to perform complete or partial code refactoring. These papers provide details about the level of customization options available in each integration solution. Overall, this paper compared each aspect that is important from the cloud migration perspective and it will help the developers to select the right integration solution for their cloud migration use case.

A suitable solution can remove the need for information technology (IT) staff to manage the infrastructure. “It has been tried to offer a solution that could run at almost unlimited scale in a pay-as-you-go model. Based on the information gathered using different sites, vendor’s white papers, web journals, correlations, and past overviews it has been observed that while migrating the legacy database to the cloud provider plays an important role as it is observed that if the legacy system is based on the Microsoft or Oracle platform, then it is always recommended that the migration should be performed on the Microsoft cloud or Oracle cloud as it provides inbuilt

**Table 1** Comparison among different integration solutions

	Azure Data Factory	AWS Glue	Informatica PowerCenter	Oracle Data Integrator	IBM DataStage
Focus	Data integration, ETL	Data integration, ETL	Data integration, ETL	Data integration, ELT	Data integration, ETL
Connects to data warehouses? Data lakes?	Yes/yes	Yes/yes	Yes/yes	Yes/yes	Yes/yes
Transparent pricing	Yes	Yes	No	Yes	Yes
G2 customer satisfaction	4.6/5	4.1/5	4.0/5	4.1/5	4.0/5
Support SLAs	Yes	Available	Available	Available	Available
Purchase process	Options for self-service or talking with sales	Options for self-service and talking with sales	Requires conversation with sales	Options for self-service and talking with sales	Options for self-service and talking with sales
Compliance, governance, and security certifications	HIPAA, GDPR, ISO 27001, others	HIPAA, GDPR	HIPAA, SOC 2, SOC 3, Privacy Shield	HIPAA, GDPR	HIPAA, GDPR
Data sharing	No	Yes, within AWS	Yes, through Informatica Cloud	Yes, within oracle cloud	Yes, within IBM cloud
Vendor lock-in		AWS Glue is strongly tied to the AWS platform. Usage is billed monthly	Annual contracts. No open source	Annual contracts. No open source	Annual contracts. No open source
Developer tools	REST API,.Net and Python SDKs		Informatica Developer Tool, REST API, Connector Toolkit		

functionalities that support the legacy application. It has also been observed that each solution has been recommended by considering the current employee strength and the revenue of the companies who wish to migrate to the cloud. As most of the insurance companies that are the core user of the healthcare analytical tools process very large size of data from different data sources so the solution should be good enough that construct ETL and ELT processes code-free in an intuitive environment. The solution should be flexible enough that the user should be able to write the customized code. Based on our study it has been concluded that Microsoft

Azure data factory or AWS (AMAZON WEB SERVICES) Glue can be the most suitable data integration solution that can be used for migrating the legacy database application to the cloud.

## 7 Conclusion and Future Scope

This paper reviewed the literature on different data integration and ETL solutions. Each tool focusing on the different issues that the clients face while performing the migrations over the cloud. This paper discussed five data integration solutions implemented in world-class organizations. Each toll has its own benefits over another, and these tools made cloud migration easy and efficient. Cloud-based data integrations solutions allow users to select the best match of UPS performance to business objectives, with an adaptable approach to changing needs. In this paper the analysis has been limited to the study of different research papers, blogs and the opinion and reviews shared by industry experts and a comparison has been performed for few industry-leading integration solutions. There is a scope for further implement the findings in practical scenarios and that will help to provide more concrete results. The practical approach has not been used because the pricing of the integration tools is very high and only using the integration tool will not be enough as the practical implementation can only be performed by creating a complete cloud architecture. Integrated software management provides actionable intelligence for an ideal balance of high availability and peak performance. The analysis has been performed from the perspective of migrating legacy healthcare analytical tools to the cloud and it has been concluded that the Azure Data factory is the best solution to perform the migration.

The Future scope involves studying other on-demand multi-tenant cloud integration platforms for connecting cloud and on-premises applications and data, cloud-based integration processes that are used to transfer data between cloud and on-premises applications. A more practical approach can be used by doing some practical implementation and by sharing more real-time experience.

## References

1. G. Tredrea et al., [https://www.researchgate.net/publication/347047830\\_CLOUDBASED\\_INTEGRATION\\_AND\\_STANDARDIZATION\\_OF\\_ADDRESS\\_DATA\\_FOR\\_DISASTER\\_MANAGEMENT\\_-\\_A\\_SOUTH\\_AFRICAN\\_CASE\\_STUDY](https://www.researchgate.net/publication/347047830_CLOUDBASED_INTEGRATION_AND_STANDARDIZATION_OF_ADDRESS_DATA_FOR_DISASTER_MANAGEMENT_-_A_SOUTH_AFRICAN_CASE_STUDY) (2020)
2. S. Westin, K. Stendal, [https://www.researchgate.net/publication/328043442\\_IMPLEMENTING\\_CLOUD\\_BASED\\_BIG\\_DATA\\_PLATFORMS\\_A\\_CASE\\_USING\\_MICROSOFT\\_AZURE](https://www.researchgate.net/publication/328043442_IMPLEMENTING_CLOUD_BASED_BIG_DATA_PLATFORMS_A_CASE_USING_MICROSOFT_AZURE) (2018)
3. M. Bodziony et al., <http://ceur-ws.org/Vol-2572/short2.pdf> (2020)
4. A. Gupta, A. Sahayadhas, [https://d1wqxtx1xzle7.cloudfront.net/66042041/IJM\\_11\\_07\\_044-with-cover-pagev2.pdf?Expires=1633664761&Signature=hAXrg-SpXyLKWS5bwB151IQI](https://d1wqxtx1xzle7.cloudfront.net/66042041/IJM_11_07_044-with-cover-pagev2.pdf?Expires=1633664761&Signature=hAXrg-SpXyLKWS5bwB151IQI)

- e8QGBVlw6YsUWIkTOFFT6XgAVbx6LL9MQSxh0grTDwaZd85To12Po2oGWWK44-aMY-czGNGZ84gwD7LAj5xLYwyux17jcIm27Hx09LvbhyFUPURbA9FXEFrll-wTwPxm hGjbFh06QgIKR5FDn-Rp5jlTqy2Jo3VbUZHZZkuPoyLuHU0hQSXUVnAgCiPvMthGCc 1qOyTCkYtB6wWXcHRvkBI-limUthl5sU0w8LIEmMLnPTmt91LaLY94J36Oll76~lIE h8KpxHiZmGiRnMbi2pDFd4Bj6gfOqbWG1G6CdIKb2yNtGdpMj4P2w1Fg\_\_&Key-Pai rId=APKAJLOHF5GGSLRBV4ZA (2020)
5. L. Wycislik et al., [https://doi.org/10.1007/978-3-319-18422-7\\_17](https://doi.org/10.1007/978-3-319-18422-7_17) (2015)
  6. Seagate, [https://www.seagate.com/files/www-content/our-story/rethinkdata/files/Rethink\\_Data\\_Report\\_2020.pdf?utm\\_source=xp&utm\\_medium=blog&utm\\_campaign=content](https://www.seagate.com/files/www-content/our-story/rethinkdata/files/Rethink_Data_Report_2020.pdf?utm_source=xp&utm_medium=blog&utm_campaign=content). Accessed 12 Sept 2021
  7. M.B. Biplob et al., <https://ieeexplore.ieee.org/document/8745574> (2018)
  8. R. Mukherjee et al., <https://ieeexplore.ieee.org/document/7976926> (2017)
  9. P.S. Diou et al., <https://ieeexplore.ieee.org/document/8277875> (2017)
  10. G. Jayashree et al., <https://ieeexplore.ieee.org/document/9132936> (2020)
  11. J. Sreemathy et al., <https://ieeexplore.ieee.org/document/9441984> (2021)
  12. U. Khan et al., [https://ieeexplore.ieee.org/document/8493901?denied="](https://ieeexplore.ieee.org/document/8493901?denied=) (2018)
  13. A. Sabtu et al., <https://ieeexplore.ieee.org/document/8002467> (2021)

# A Proposed Structure, Primarily Based Totally on Cloud Computing for Enhancing Electronic Commerce Applications



Neha Tiwari and Navneet Sharma

**Abstract** The cloud companies have caught the IT Company's attention, to give new approach of services concept. This is particularly seen in the companies of e-commerce. At present, E-commerce companies and associations, has a major environmental cost problem because of with the cloud computing advent, the aim of this paper is to introduce the suggested applications of ecommerce and their infrastructure based on various concepts and origins, cloud computing development trends on the basis of resource storage and e-commerce related problems. With the help of proper implementation the cost for the activities of ecommerce and problem solution at company level is insufficient to make proper activities regarding e-commerce due to unavailability of sources.

**Keywords** Cloud computing · E-commerce cloud · SaaS · IaaS · B2B · PaaS · B2C

## 1 Introduction

One of the most popular and well known growing applications on web is E-commerce, these app allows the shoppers and customers and employees to get easy and fast access to shopping and services. According to general e-commerce definition as presented by the Association of Electronic Commerce, the business of any form is covered by electronic commerce in the form of business related transactions made at the time of using data or communication technology. Depending on the transitions and their nature, e-commerce is distinguished in following main types, business to business, business to customer, customer-to-customer, customer to business, e-commerce with

---

N. Tiwari (✉) · N. Sharma  
IIS Deemed to be University, Jaipur, India  
e-mail: [neha.08sep@gmail.com](mailto:neha.08sep@gmail.com)

N. Sharma  
e-mail: [navneet.sharma@iisuniv.ac.in](mailto:navneet.sharma@iisuniv.ac.in)

no business, with the internet access by non-business related organizations like educational and government institutions to cut down the cost and enhance their services and also e-commerce within a business.

All important business related data is offered by system of ecommerce that include the cost of merchandise and marketing quantities and in this way, it helps in the facilitation of various commercial executions. When the technological information and knowledge is used at the time of space has resulted in the elementary variations inside the method and all these commercial activities are considered particularly important. These commercial activities can result increased in the pricing dynamically and hence there is a need of comparisons at merchandise level and more frequent barter contracts take place. There are some highly advanced and modernized techniques related to IP, these techniques are used to deliver the bearing, and hence, there will be higher potential for more changes [1].

The fashions regarding E-fashion are nevertheless inside the exploration degrees of early stages and other applications, some issues are always there that are linked with the fashions such as security of the platform and technological standards, the different regulatory offerings and these are not resolved naturally but they are kept in practice. On the other hand, the cloud computing appearance can be a cure to many such issues.

One of the foremost booming technologies is known as cloud computing and it is among the data technology skills and business in addition since inside the house the physical properties are considered and they account for the higher support of the code and hence they keep the infrastructure protected. The additional technology specialists are attracted towards it as well. In the good economy, cloud plays a significant role and the regulative changes and its potential is needed for the better implementation of various applications by the cloud computing potential victimization [2].

The variety of computing is the name given to cloud computing in most of the cases especially when there is a connection of massive ascendable information and potential. The service is provided across the internet so that multiple external customers can take benefit from it [3]. From the distributed computing system, it associates evolution comprising of the computers that are interconnected and assorted offering dynamic services along with other computing and unified resources supported by various agreements at service level.

Cloud computing is one of the most modern technology and this is one of the main advantageous aspects and lower cost is needed for the implementation of its general infrastructure such as license, hardware, code etc. There are certain high profile businesses such as Google and Microsoft supplying the cloud at no cost for the main academic units and educational systems. They also offer free of cost cloud to the business related applications. The cloud computing as a fresh and modern model of service along with higher network storage potential, on demand nature access offers the ideal alternate to the sharing of resources and process mechanisms. Inside the conditions existing, the frameworks of cloud computing allow less investment from enterprises to applications of e-commerce (B2B and B2C). At present, e-commerce mixture and analysis of cloud computing focuses at the technological level and

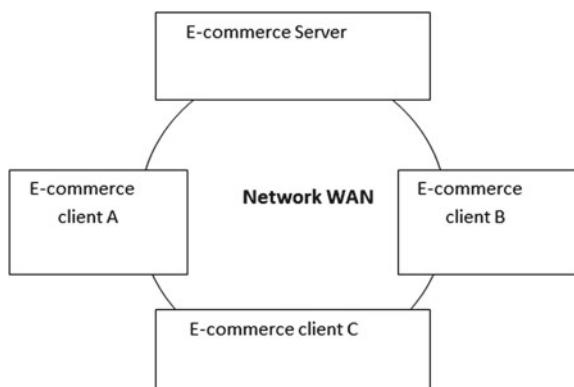
hence, e-commerce applications are considered generally in the cloud computing for the servers and other communication frameworks [4–6] (Fig. 1).

The consumer hardware may well be a mobile device or a desktop computer. The client application will be an easy application or an avid application. The e-commerce server can use cloud computing, therefore all the desired resources are going to be adjusted PRN [4].

E-commerce systems can use like cloud computing victimization [8]:

- The main infrastructure, on the infrastructure of provider, uses an e-commerce resolution.
- Platform: the developed e-commerce solutions are used and these are often supported by the development of interface of the provider.
- Products and services: the e-commerce solutions are presented by the provider. There is a wider acceptance for the cloud computing these days and its one of the most beneficial aspects of cloud computing [4]:
- Low price is seen, there is no price conjointly or sometimes for the update of hardware only terribly tiny ones are used.
- Adaptability and flexibility: to maximize the investments, the infrastructure may be scaled. As the demands often seen fluctuating, the cloud computing allows dynamic measurability.
- Accessibility: the knowledge creation is facilitated by it and other public services can be obtained by it at the same time, no sensitive problems are made vulnerable.
- There are certain gadgets with necessities of hardware such as cell phones whereby the cloud is used in the form of cloud consumers.
- There is no need to transfer or install any specific software solely to become a part of cloud and there is no need to associate it on the net.
- There is no need for the crash recovery in case of this cloud system. If the computer of a client crashes, all the information is lost. But in cloud computing method, all the data is stored inside the cloud. In addition to it, there are some other advantages of cloud computing system as well [8].

**Fig. 1** E-commerce framework [7]



There are some disadvantages as well in the cloud computing system: these include,

1. The overall performance is impacted by the net connection speed.
2. Over a long haul, the subscription fee for the information center is also more than hardware shopping.
3. There is a crucial need of service quality and backups are needed as they are essential to keep the data secured.

There are some major security advantages as offered by the cloud computing services for the companies as well as people using the service and it offers ideal solution for the e-commerce and there are following benefits of using this technology

- Improbability is increased
- Knowledge access is observed more
- The data storage is generally centralized
- There is more virtualization in it.

## 2 Related Work

The significant attention has been drawn by the system of cloud computing from investigators in past few years. Wang [9] for example, mentioned and highlighted a system that has a potential to increase the cloud computing environment and improve the ecommerce system. For this purpose, It established a latest and modern model for the exploration of some problems. He parsed the connection related concerns in the model for the development of e-commerce and cloud computing was added as well to support it. The essential plan of building new and modern business based models were the main aim of this investigation on the basis of cloud computing. He also explored the issues related to various concerns since the resolution acceptable for the technology is offered by cloud computing. The level of exploration is revealed and along with it, there is also mentioned some ways which can be followed to use the cloud computing technology ideally and to open up the projects of e-commerce and development of some highly needed patterns for replacement of this technology.

Another investigative worker, Liu [10] introduced the model for the applications of e-commerce and also tried to cope with the demands and other concerns regarding e-commerce. He also mentioned the lack of some resources and this was all done by the establishment of the ecommerce applications in a proper framework on the basis of cloud computing service models.

Liu worked on the resolution of some concerns regarding funds shortage, the use of labor or man power and other technological aspects which he considered are needed for the e-commerce related activities. He also demonstrated that in all the companies and top enterprises, the network's resources are used for e-commerce. E-commerce applications that are based on cloud computing allow the lower prices to enterprises and this is done by carefully implementing the activities regarding e-commerce and solving of some matters at the company level and even this can

be done in the absence of resources. This makes cloud computing exceptionally important for the companies.

Another research worker Jie [11] worked in the dimension to resolve the security of group applications in C2C E-commerce. On the basis of cloud computing, he generated a visualization model. The algorithmic programs were mistreated by the trust cloud merger and hence there is a trust cloud computing issues in the cloud computing and sub attributive evaluations are there for the achievement of comprehensive attributes. The 2 facet transaction in the line of evaluation got some highly strategic advises and it has been shown in the experiments that evaluation method was valid and practical in most of the cases.

Not only to C2C in e-commerce, we used this method, but this method was rather used as one of the most valuable approaches for the trust value in the networks that are kept opened.

Two other investigative workers named, Sun and Yuan [12] developed on online syndicate devices for online selling with the framework of spring. In this various apache were used as initiatives to supply the market place generally on the basis of net service versions of software depending on the requirement of customers. There are numerous channels and tenants to support this program with the help of mode changes in the controller cluster model and these are designed especially for the sake of decentralization for the proper manipulation of cloud computing.

Take a speedy look at the specially designed algorithms for cloud computing and are designed especially for the sake of green and natural manipulation and a special kind of burst at Armbrust et al. [13]. A complete evacuation of cloud computing is needed for the ideal furnishing of it. The possibilities are considered to enable its smooth functioning and other boundaries are determined so that the proper adaptation is obtained. The providers of cloud servers such as Juncai and Shao [14] have already mentioned that era of utilizing cloud is inside the field of commercial enterprise.

The improvement is shown by their delivery in the system of cloud computing and all the concerns regarding environmental problems and challenges. The e-commerce is the most modern form of business of modern era. In the same way, the safety issues of the community and possible solutions often estimate the cloud and its related traits.

### 3 E-commerce Structure Based on Cloud Computing

The cloud computing system is based on the framework of e-commerce and there are five layers in which this system is divided. (1) Layer of hardware resource, (2) layer of software package, (3) layer of resource management, (4) server layer and 5th one is the business layer. It is shown in the figure that the framework needs proper design. The sections mentioned subsequently are some of the key elements that are suggested and some additional details are included as well.

### ***3.1 Hardware Resource Layer***

This is regarded as the inner and lowest layer inside the middle ware of the cloud computing services. For the complete infrastructure the entire framework is vital and with the help of this the system will run smoothly. To supply versatile accommodative platform the physical hardware layer is needed to enhance the utilization of resources. There are two layers that serve as a key to modern enterprise centers of information whereby the complete services are provided. The virtualization of atmosphere and the layer of management are these two important layers and data center records are stored inside these layers and managed expeditiously and can be rapidly designed and provisioned as well. When virtualization is given there are a lot of physical servers, classifiers and storage networks that are referred as platform for the package of software updates. For the uninterrupted power supply, the servers of cloud middleware are there for the e-commerce systems that are based primarily on the cloud. There is a dynamic expansion of the physical host in the pool and ascendable memory is seen as well and in this way some additional memory related features are considered in this system.

### ***3.2 Software Resource Layer***

There is a formation of subsequent second layer, it can address the operational concerns, and the technology used is the middleware technology. There are several codes of computers that are added for the interface to be sorted for the developers of software interface. Hence, the developers of software are likely to create the e-commerce based applications and are mostly ready to take part in the cloud space enabling the users of cloud to cyber such apps with the help of services of cloud.

### ***3.3 Resource Management Layer***

In the last two layers and their management, this is considered as an important layer. The computerized codes are lost in this due to coupling with some other hardware resources. The programming plan and virtualization assistance are used and uninterrupted demands are obtained for various resources of hardware management.

### ***3.4 Service Layer***

There are three sublevels whereby we have divided these layers and these are named as IAAS, PAAS and SAAS. The cloud customers are facilitated by the

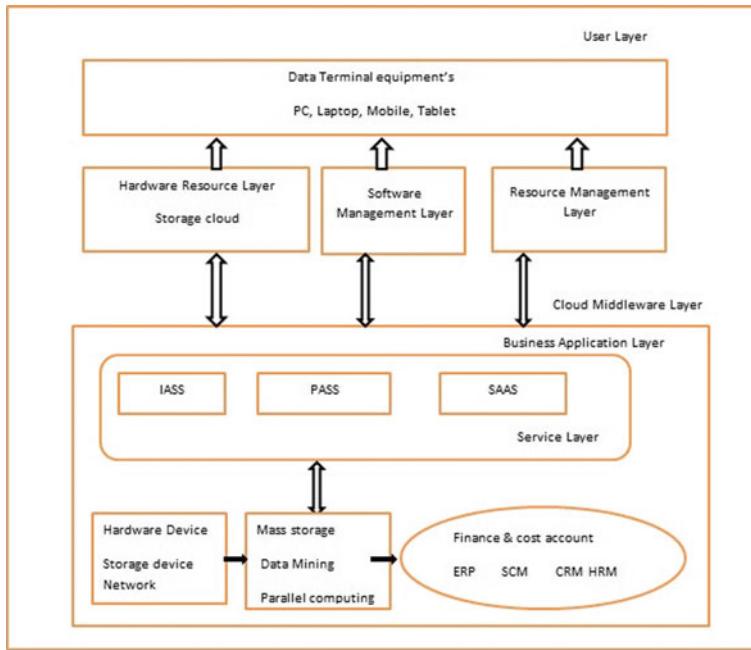


Fig. 2 E-commerce system based on cloud computing [7]

cloud customers for the package resource products, and some other resources also involved in this service layer such as resources regarding hardware and framework or infrastructure (Fig. 2).

### 3.4.1 Infrastructure as a Service of E-commerce Cloud

In the cloud space that is used in e-commerce, the digital users of cloud space can get easy and fast access to the digital service at any time. The required network is acquired by the users and natural computing services are available at any location. The house and physical spaces are integrated and in this way potential of computing increases as well. This leads to the better integration of digital and cloud services and information terminals along with the instrumental system and in the future, e-commerce will be based on such modern features.

### 3.4.2 Software as a Service of E-commerce Cloud

The applications package package or services provided by knowledgeable company e-commerce, the companies to pay among the similar approach of on-demand access, in line with the number to calculate the worth, complete the assembly, marketing,

mercantilism and management. Companies use ecommerce system in lower value to avoid wasting, and plenty of resources could also be used for business activities. E-commerce cloud setting provides useroriented ubiquitous adjective hardware resources, computing atmosphere and package services.

In e-commerce cloud area, users can access to digital services transparently at any time in anyplace. The users will acquire the specified network and computing services really naturally at any position. The information house and physical area square measure progressing to be integrated as a result of ubiquitous computing capability. And additionally the everpresent info terminals beside the embedded system instrumentation are the vehicles of ecommerce among the long run.

### ***3.5 Business Application Layer***

From all the alternative layers, the business application layers in the cloud computing system is based on the architecture of ecommerce and the virtual e-commerce layer to work efficiently. There are some additional frameworks of enlargement based on the e-commerce clusters and their parts. The SAAS services are gaining more popularity and there are various aspects of potential utilization of such services as they are supposed to increase the cloud platform leverage and run the business related processes in much efficient way. The business processes that can be run in ideal way include the management of chase, creation of merchandise, estimation, management of client service etc. In the same way, client service portals, chain management provision are also included along with accounting of valuable finances. The human resource management and enterprise resource management play efficient role in it.

Therefore, in USA the cloud based system is maintained in such a way that maintenance, purchase and IT workers training like aspects are based on the cloud system. They work efficiently with the system code of the software. Inside the cloud, one network's technology answers and in this way the purchasing operations are made. There is lesser maintenance time and less replacement of tools and less expensive equipment's are needed for its up gradation. The industrial leading time is allowed by the cloud solution conjointly to time it up with price and in this way the time of implementation averages below the ancient supply chain solutions of the time 0.5. Along with this, across all the value chain participants, the real time knowledge means more opportunities for review in the future. In this way, there is a use of some costly tools and programming languages are available for speedy and easy maintenance of this system. The pricing network is checked and familiar instant data flow is estimated as well. The value network frame is considered to hook up with entire value at the cloud services and in this way you can reduce the cost and increase the earning.

## 4 The Importance of Unification Between Cloud Computing and Ecommerce

With the development of EC and its improvement, there are some hard and fast rules and there is always a need to convert various constraints due to the company size concerns, economic potential of a company and technological force that is confirmed in particular in some of the aspects mentioned below [15, 16].

1. The mass statistics garage is included in technical issues such as mining stats and protection of information. We need to come back hard and to maintain such aspects for the ebusiness of small and medium sizes.
2. The gadgets and operational cost are confirmed to make it sure that there is a considerable improvement in the system of e-trade.
3. There are various limitations in the capability processing and overall performance protection. There are certain imperfections hindering the e-trade improvement. In none of the cases, the cloud computing can be useful in the development of other systems for e-trade and there is always a possibility to expand business. This can be confirmed as mentioned below:
4. The cloud computing serves as a basis of e-commerce and etrade and in the same way massive scale is there surrounding the middle stats along with the computation at higher pace and statistic garage. The mining potential of the statistics are considered and commercial enterprises are made to benefit and expand the businesses working on e-trade pattern.
5. The cloud models in case of e-trade are used in the distributed pattern in the pattern. The centralization is controlled by the process of virtualization which is considered to be the basis of all the statistical data.
6. Since there is a chance of entry of statistics in the fairly centralized manner, the virtualization process is easier and fast in processing and it can transmit the data and keep it secured.
7. The cloud computing based operations and primary controls are kept under check. The commercial E-trade is used to check the flexibility of statistics and operational price limitations. The software related answers are estimated without even considering the resource and location of any particular gadget [17].
8. SCM is an abbreviation of supply chain management and it is based totally on the system of cloud computing. There is a steady and reliable carrier of statistical garage and customer requirements are calculated at each time [17].

## 5 The Primary Problems of Ecommerce Cloud

In any e-commerce based model that is made newly, the cloud computing can be added. The cloud computing and e-commerce are mixed and they can attain a considerable mature stage and can be practiced well. Currently, several issues are there and they are needed to be solved in first priority [15, 16, 18].

### ***5.1 Security Issues of Cloud Computing***

There is always a need of privacy and confidentiality in the cloud system and its security. At present, in the cloud computing of every kind, there is a mixed, public and nonpublic cloud and all these concepts are planned efficiently. They are used in common by the businesses and additional verification of such resources is needed.

### ***5.2 Defiance of Cloud Applications***

For the companies working on the mechanism of e-commerce, there is a need of components at risk while third party is working. The risks and benefits are estimated and hence, to hook up with the third party, you will be troubled to develop such systems.

### ***5.3 The Standards of Cloud Computing***

The cloud-based e-service model continues to be during a fragmented state. If the users really need to market and apply these new models, a unified business normal ought to be developed.

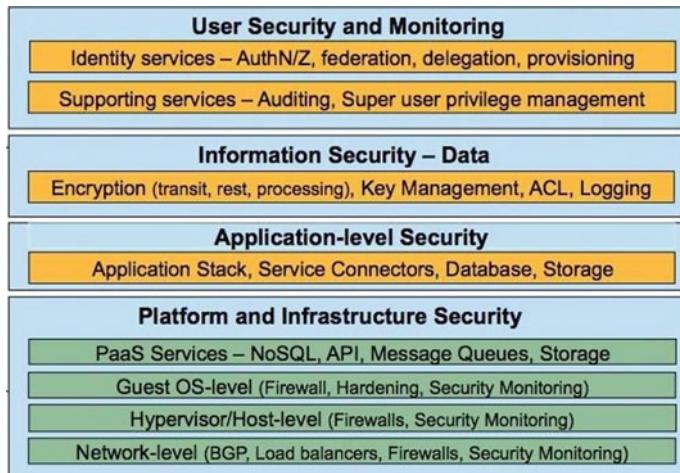
### ***5.4 Regulatory Issues of Cloud Services***

There are some products and services that are supposed to be given to the platform of cloud computing. In the information storage processing, data processing, the management of security and its maintenance, the cloud services are used typically in the environment of cloud computing. Hence the position of the service provider is always important to ensure that information is kept secured and all the data is kept hidden and confidential hence, to monitor the service and scalability, there are some important issues to be resolved.

## **6 Suggested Layered Framework**

Figure 3 provides the layer, provider and user responsibilities within a cloud service that are secured by the provider versus the customer.

We suggest an enhancement that whenever the user signing up with a provider, it is important to read a gap analysis on the cloud service capabilities. For this practice



**Fig. 3** Layer framework [19]

providers should give them an “Architect for security-as-a-service”, which would benchmark the cloud platform’s maturity, transparency, compliance with enterprise security standards and regulatory standards. This would be the best practices to mitigate risks to cloud services.

## 7 Conclusion

In the present time, in the environmental pricing there is a particular drawback and the e-commerce makes use of all such aspects. The cloud computing therefore serves as an ideal platform to resolve such issues in modern businesses and applications of e-commerce. The framework of ecommerce is supported on the basis of computing at cloud and this is how the text is introduced. These models of ecommerce are still in the explorations to be used in various applications. The network security is always the key and there is a need to evaluate the security of network to keep the data confidential otherwise e-commerce applications would stop working and they will show a major flaw in their execution. There is a need to refine cloud technologies additionally for the cloud computing applications before they are more distributed and widespread. A brand new era of cloud computing is therefore inaugurated by the ecommerce. For this purpose, we proposed a structure of cloud computing standards of security to minimize the human errors and improve operational efficiency of Ecommerce applications.

## References

1. M.J. Shaw, Electronic commerce: state of art, in *Handbook on Electronic Commerce*, chap. 1, ed. by M. Shaw, R. Blanning, T. Strader, A. Whinston (Springer, 2001), pp. 3–24
2. D. Bcigalupo, G. Wills, D. De Rouse, V. Chang, A categorization of cloud computing business models (IEEE/ACM, 2010)
3. M. Head, K. Hassanein, Trust in ecommerce: evaluating the impact of third-party seals. *Q. J. Ecommerce* (2002)
4. P. Paul, A. Felician, V. Marius, Measuring the efficiency of cloud computing for e-learning systems. *WSEAS Trans. Comput.* **9**(1) (2010)
5. Y. Wang, K.-J. Lin, Reputation Oriented Trustworthy Computing in E-commerce Environments (2007)
6. C. Boja, L. Batagan, Software characteristics of M-learning application, in *WSEAS International Conference on Mathematics*, 23–29 March 2009
7. A.A.E. Saleh, A proposed framework based on cloud computing for enhancing e-commerce applications. *Int. J. Comput. Appl.* (0975 – 8887) **59**(5), (2012)
8. C.L. Li, Z.H. Deng, On the value of cloud computing. *Library Inf.* **4** (2009)
9. X. Wang, Research on e-commerce development model in the cloud computing environment, in *International Conference on System Science and Engineering (ICSSE)*, 30 June 2012–2 July 2012
10. T. Liu, E-commerce application model based on cloud computing, in *International Conference on Information Technology, Computer Engineering and Management Sciences (ICM)*, 24–25 Sept 2011
11. Z. Jie, Trust evaluation model based on cloud model for C2C electronic commerce, in *International Conference on Computer Application and System Modeling (ICCASM)*, 22–24 Oct 2010
12. J. Sun, C. Yuan, A multichannel online selling syndicates based on cloud structure, in *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, 2012
13. M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica et al., Above the Clouds: A Berkeley View of Cloud Computing. Tech. Rep. UCB/EECS-2009-28, EECS Department, University of California, Berkeley, 2009
14. S. Juncai, Q. Shao, Based on cloud computing e-commerce models and its security. *Int. J. e-Education e-Business e-Management e-Learning* **1**(2) (2011)
15. K. Gunasekar, C. Anirudh, Analysis of security issues in cloud based e-learning. Master's thesis in informatics, University of Boras, Aug 2011
16. T. Liu, E-commerce application model based on cloud computing, in *International Conference on Information Technology, Computer Engineering and Management Sciences (IEEE, 2011)*
17. M. Tim, K. Subra, L. Shahed, *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance* (O'Reilly Media, Inc., 2009)
18. S. Dan, C. Roger, Privacy and consumer risks in cloud computing. *Comput. Law Secur. Rev.* **26**, 391–397 (2010)
19. S. Kumaraswamy, infoQ Newsletter. Introduction to Cloud Security Architecture from a Cloud Consumer's Perspective, Dec 07, 2011

# Deep Learning and Internet of Things (IOT) Based Irrigation System for Cultivation of Paddy Crop



S. Sasikumar, D. Naveen Raju, P. V. Gopirajan, K. Sureshkumar, and R. Pradeep

**Abstract** Agriculture has a significant role in cultural life, and Agriculture has a significant role in the economies of many countries. To come up with the best possible conclusion from this study. It's essential to pay attention to critical factors, including energy, water availability, labour, and a correct watering plan for crops. Researchers in this study were interested in building a smartphone application that would allow farmers to operate an IoT-based automated irrigation system remotely. Paddy field photos were used to create a deep learning model called Paddy Field Radial Basis Function Networks (PF-RBFNs). The model tells the farmer how much water will be needed in a certain field area for irrigation. A real-time picture dataset and a raspberry pi-based hardware model were used to test this approach. The model was compared to three different deep learning models: LSTMs, RNNs, and GANs. It was found that this proposed PF-RBFNs model has a 93% accuracy rate.

**Keywords** Radial basis function networks (PF-RBFNs) · LSTMs · RNNs · GANs

## 1 Introduction

The main motive of this project is to solve the problem of farmers related to irrigation (water) because a farmer is a magician who produces money from the mud. Agriculture is one of the oldest and prime cultures of human beings [1]. It is the backbone of the Indian economy. Though the technology in agriculture is developing, the transfer of technology to real-time implementation is a challenging task [2]. The main reason

---

S. Sasikumar · P. V. Gopirajan (✉) · R. Pradeep  
Department of CSE, Saveetha Engineering College, Chennai, India  
e-mail: [gopivrajan@gmail.com](mailto:gopivrajan@gmail.com)

D. Naveen Raju  
Department of CSE, Sri Sairam Institute of Technology, Chennai, India

K. Sureshkumar  
Department of Information Technology, Saveetha Engineering College, Chennai, India

for the loss in agriculture is the “low Production of crops”. Water scarcity and waterlogging have a huge impact on the low production of crops. Water scarcity is a lack of water to irrigate the land, and waterlogging is nothing but plants die due to excessive water [3].

Many approaches were addressed to the similar issues faced by the farmers. The authors suggested many recent advancements in IoT implementation for improving the smart agriculture process [4]. Geographic Information System (GIS) based models were developed with the implementation of IoT suggested a natural and artificial landscape [5]. Using traditional methods to dispose of agricultural wastes produces pollution and damage to the environment [6]. Implementation of artificial intelligence and machine learning, irrespective of the domains, helps achieve good results in predicting yield [7].

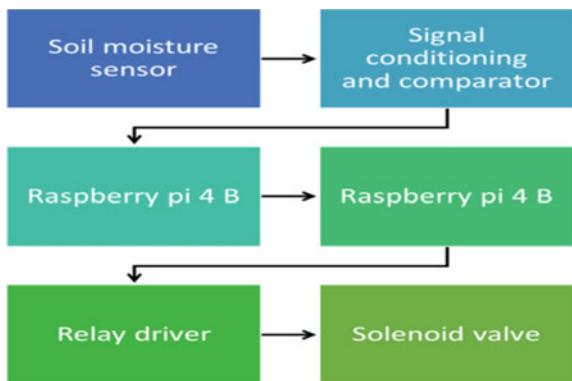
Deep learning models provide high accuracy [8–10]. Deep learning algorithms such as Convolutional Neural Networks (CNNs) [11–13], Long Short Term Memory Networks (LSTMs) [13], Recurrent Neural Networks (RNNs) [14–16], Generative Adversarial Networks (GANs) [17–19], Radial Basis Function Networks (RBFNs) [20–22], Multilayer Perceptron (MLPs) [23–25], Self-Organizing Maps (SOMs) [26, 27], Deep Belief Networks (DBNs) [9, 28–30] and Restricted Boltzmann Machines (RBMs) [29, 31–33] were widely used in many IoT related applications.

To solve the irrigation problem and save water, this study proposed better technology to help the farmers implement IoT and deep learning models, named “BIOT IRRIGATION SYSTEM”. This system deals with the automatic water sprinkler system, which senses itself. The moisture content of the soil will be notified along with the images captured through a drone camera and predicts the amount of water required. Finally, this BIOT system suggests the prediction of whether farmers need to irrigate or not. It also suggests how much water is needed for the soil. The farmers can control through an app connected with BIOS IRRIGATION SYSTEM through the data stored in the cloud. Validation of this proposed study was carried out by comparing similar algorithms, and their efficiency was recorded.

## 2 Methodology

The main objective of this project was to provide adequate water to the crops, which can save water. Farmers can automatically water the plants with simple clicks in the mobile application. By this, the farmers minimize the land workers. An automatic plant watering system using a microcontroller gives the interrupt signals to the motor through a relay that is programmed already. Arduino board is connected with a soil moisture sensor that senses the moisture content, and the pump motor is activated by giving a signal to the microcontroller. Automatic irrigation systems are suitable for those who move from place to place. Thus the “Deep learning and IoT based Automated Irrigation System was developed, and working was carried out based on soil moisture using raspberry” has been designed with deep learning algorithm and tested successfully.

**Fig. 1** Proposed system architecture



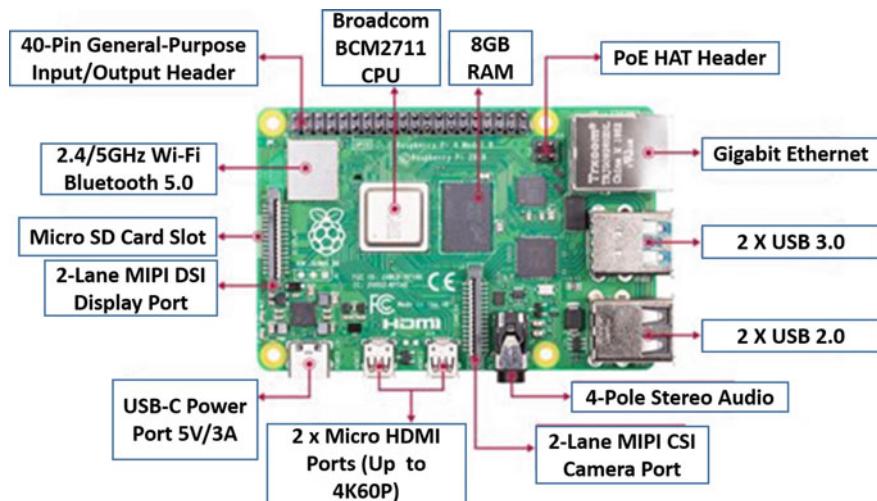
The BIOT system provides an app using Android app development (Java/XML). By using this app, a farmer can irrigate his land whenever he needs it. In this app, we place a list of plants that appears on his screen. The reason for providing this list is each plant has different characteristics (moisture level). In this list of plants, a farmer has to click a crop that he wants to Irrigate. Once clicked, information has been sent to the IoT system (BIOT) through a cloud. The BIOT system is nothing but a robot placed with raspberry pi 4 B connected to the moisture sensor. The reason for the connection is each plant has different moisture levels. The robot will check each plant's moisture level. If the moisture level is normal, it doesn't need water. If the moisture level is low, the signal will be sent to another raspberry pi 4B, placed with a solenoid valve to provide water for that particular plant. This can be programmed with the quantity of water level needed for that particular crop.

By the way, the entire crop will be monitored. After completion of irrigation, it will send a message to the user app. Figure 1 shows the proposed system architecture where the soil moisture sensor provides the input signal to the signal conditioning and comparator module. The Signal conditioning and comparator module compares the received input signal and removes the noise. A clear and processed signal was given as the input to the raspberry module. Relay driver and solenoid valve suggest the required information to the user via mobile application.

## 2.1 *Module/Component Description*

### 2.1.1 **Raspberry Pi**

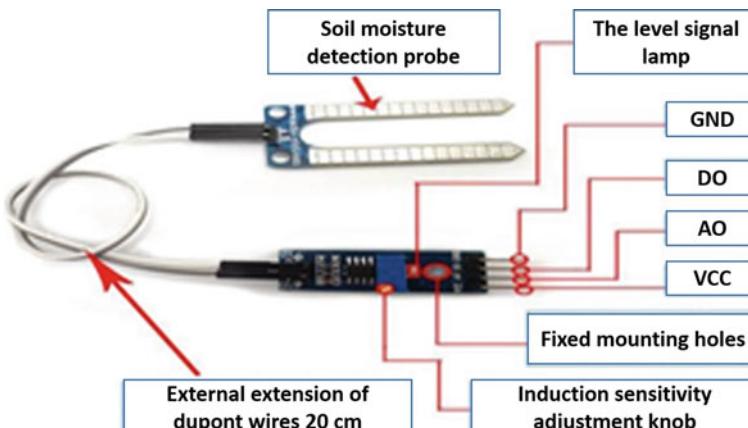
We may use the Raspberry pi's GPIO (General Purpose Input/Output) pins to control electrical devices for physical computing and experiment with the internet of things (IoT) related algorithms. The 8 GB Random Access Memory (RAM) on this raspberry module is depicted in Fig. 2 for quicker processing.



**Fig. 2** Raspberry Pi module

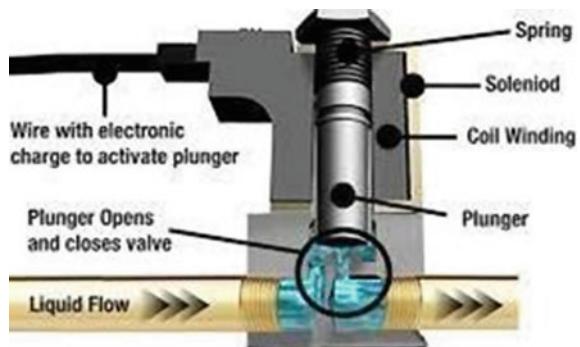
### 2.1.2 Moisture Sensor

The volumetric water content of the soil is measured and estimated using the Soil Moisture Sensor depicted in Fig. 3. Since the removal, drying, and weighing of a sample is required for direct gravimetric determination of free-soil moisture. The RED light on the comparator chip indicates that the module's output is high; thus, the probe is dipped into the dirt. That way, water pours from the water pump servo motor.



**Fig. 3** Soil moisture sensor

**Fig. 4** Parts of a solenoid valve



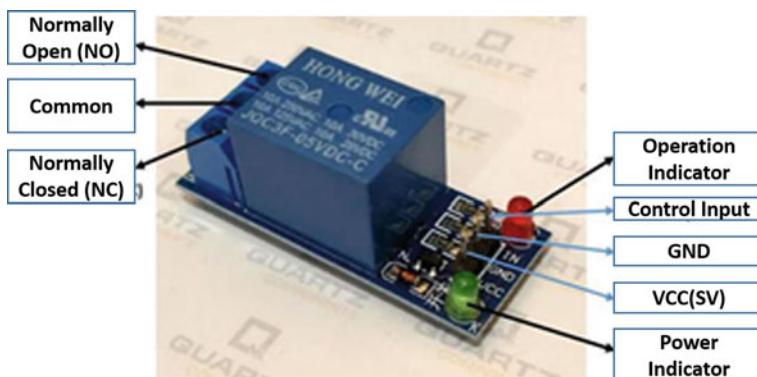
### 2.1.3 Solenoid Valve

Figure 4 depicts the Solenoid valve's interior components. An electromagnet is a coil of wire used to generate a magnetic field, and the word “solenoid valve” is used to describe any device that uses a solenoid to generate mechanical energy. To generate linear motion, an electric current generates a magnetic field.

### 2.1.4 Relay

As shown in Fig. 5, this relay is an electromagnetic switch that allows circuits to be switched by electrical equipment. The relay can be connected to the raspberry pi 4 B, which gives the signal, and another side of the relay is connected to a servo motor for passing current to it. Thus, the servo motor gets a signal from the relay.

All the components are finally assembled, and raspberry pi 4B(1) was programmed with how much moisture the plant of different species has. Then the sensed data can



**Fig. 5** Relay

be sent to the raspberry pi 4B(1), which decides whether watering is needed or not. If needed, the signal to send to the next raspberry pi 4B(2) is programmed with how much water is needed for the specific plants. Then the signal pass through the relay, the relay will instruct the servo motor to pour water into the plants.

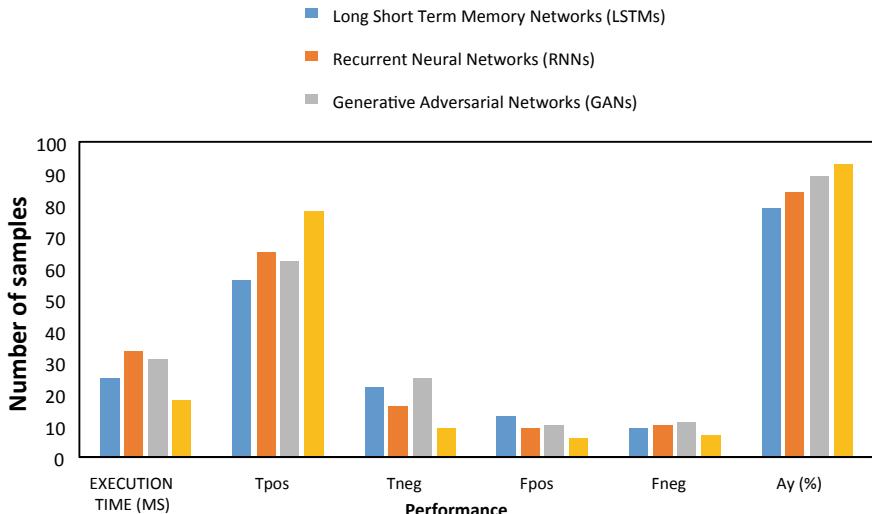
### 3 Results and Discussion

This proposed model was implemented in a real-time environment. An irrigation system was implemented and studied at the Paddy field situated at Padalam village, Chengalpattu District, Tamil Nadu, in Fig. 6. The developed Paddy Field-Radial Basis Function Networks (PF-RBFNs) model was validated with the implementation of the same dataset with other similar algorithms, namely Long Short Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs). Proposed PF-RBFN was compared based on the True Positive (TPos), True Negative (TNeg), False Positive (FPos), False Negative (FNeg), Accuracy (Ay), Precision (Pn), and Recall (Rl) parameters.

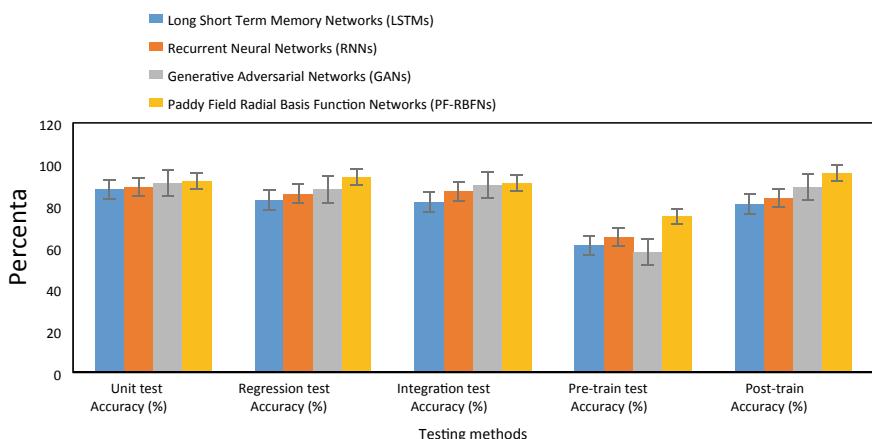
Figure 7 compares the proposed PF-RBFN model and the other relevant deep learning models—LSTM, RNN and GAN. PF-RBFN model achieved 93% accuracy when compared with other models LSTM (79%), RNN (84%) and GAN (89%). Additionally, the execution time was calculated for the PF-RBFN model and the other models like LSTM, RNN and GAN. PF-RBFN model achieved 18 ms, LSTM got 23 ms, RNN got 33 ms, and GAN got 30 ms. This shows that the proposed PF-RBFN achieved good and less execution time, and the accuracy was high.



**Fig. 6** Farmland



**Fig. 7** Performance comparison



Similarly, the proposed model and other comparative models were tested with regular testing methods such as unit testing, regression testing, integration testing, pre-train test and post train test. Proposed PF-RBFN attained more accuracy % compared with the other models. This shows that the proposed PF-RBFN model works well in predicting the amount of water required for irrigation with the BIOT system.

The overall process of the BIOT was attempted and made successful. This system seems to give that the more accurate result. The accuracy of the BIOT was 93% which is the more accurate and successful result of this project.

## 4 Conclusion

Thus the “BIOT IRRIGATING SYSTEM” has been designed and tested successfully. All the hardware components are used, and the presence of every module has been reasoned out and placed successfully. The soil moisture sensor senses the moisture level of each plant. If the moisture level is below the normal moisture level, then the raspberry pi sends the signal to the relay then the relay is instructed to open the servo motor to pour water. The special thing about this project is that it will automatically check each plant and work without human resources. Thus the functionality of the entire system has been tested thoroughly, and it is said to function successfully.

## References

1. Y. Li, Y. Zheng, Regional agricultural industry economic development based on embedded system and Internet of Things. *Microprocessor Microsyst.* **82**, 103852 (2021)
2. U. Sanga, H. Park, C.H. Wagner, S.H. Shah, A. Ligmann-Zielinska, How do farmers adapt to agricultural risks in northern India? An agent-based exploration of alternate theories of decision-making. *J. Environ. Manag.* **298**, 113353 (2021)
3. T.N. Vasa, S. Pothanamkandathil Chacko, Recovery of struvite from wastewaters as an eco-friendly fertilizer: review of the art and perspective for a sustainable agriculture practice in India. *Sustain. Energy Technol. Assess.* **48**, 101573 (2021)
4. B.B. Sinha, R. Dhanalakshmi, Recent advancements and challenges of Internet of Things in smart agriculture: a survey. *Futur. Gener. Comput. Syst.* **126**, 169–184 (2022)
5. L. Kang, Street architecture landscape design based on Wireless Internet of Things and GIS system. *Microprocessor Microsyst.* **80**, 103362 (2021)
6. S. Khanal et al., An episode of transboundary air pollution in the central Himalayas during agricultural residue burning season in North India. *Atmos. Pollut. Res.* 101270 (2021)
7. P.V. Gopirajan, K.P. Gopinath, G. Sivaranjani, J. Arun, Optimization of hydrothermal gasification process through machine learning approach: Experimental conditions, product yield and pollution. *J. Clean. Prod.* **306**, 127302 (2021)
8. M. Toğaçar, Z. Cömert, B. Ergen, Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks. *Chaos Solitons Fractals* **144**, 110714 (2021)
9. P. Heus et al., Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. *BMC Med.* **16**(1) (2018)
10. Y. Yang, M. He, L. Li, Power consumption estimation for mask image projection stereolithography additive manufacturing using machine learning based approach. *J. Clean. Prod.* **251**, 119710 (2020)
11. F. Huang, X. Qi, C. Li, W. Hu, Aerial image classification by learning quality aware spatial pyramid model. *Future Gener. Comput. Syst.* **111**, 271–277 (2020)
12. X. Wang, C. Wang, Time series data cleaning: a survey. *IEEE Access* **8**, 1866–1881 (2020)
13. A. Braun, A. Borrman, Combining inverse photogrammetry and BIM for automated labeling of construction site images for machine learning. *Autom. Constr.* **106**, 102879 (2019)
14. V. Sagan et al., Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **205**, 103187 (2020)
15. W.-J. Yu, Z.-D. Chen, X. Luo, W. Liu, X.-S. Xu, DELTA: a deep dual-stream network for multi-label image classification. *Pattern Recognit.* **91**, 322–331 (2019)
16. M. Tan, J. Zhou, Z. Peng, J. Yu, F. Tang, Fine-grained image classification with factorized deep user click feature. *Inf. Process. Manag.* **57**(3), 102186 (2020)

17. A. Lumini, L. Nanni, Fair comparison of skin detection approaches on publicly available datasets. *Expert Syst. Appl.* **160**, 113677 (2020)
18. W. Zhang, Y. Zhu, Q. Fu, Semi-supervised deep transfer learning-based on adversarial feature learning for label limited SAR target recognition. *IEEE Access* **7**, 152412–152420 (2019)
19. G. Yang, F. Ding, Associative memory optimized method on deep neural networks for image classification. *Inf. Sci. (Ny)* (2020)
20. J. Schmidt, M.R.G. Marques, S. Botti, M.A.L. Marques, Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **5**(1) (2019)
21. B. Sajjadi, A.A.A. Raman, R. Parthasarathy, S. Shamshirband, Sensitivity analysis of catalyzed-transesterification as a renewable and sustainable energy production system by adaptive neuro-fuzzy methodology. *J. Taiwan Inst. Chem. Eng.* **64**, 47–58 (2016)
22. S.F. Ardabili, B. Najafi, M. Alizamir, A. Mosavi, S. Shamshirband, T. Rabczuk, Using SVM-RSM and ELM-RSM approaches for optimizing the production process of methyl and ethyl esters. *Energies* **11**(11) (2018)
23. D. Petković, Ž Čojbašić, V. Nikolić, Adaptive neuro-fuzzy approach for wind turbine power coefficient estimation. *Renew. Sustain. Energy Rev.* **28**, 191–195 (2013)
24. J.K. Whiteman, E.B. Gueguim Kana, Comparative assessment of the artificial neural network and response surface modelling efficiencies for biohydrogen production on sugar cane molasses. *Bioenergy Res.* **7**(1), 295–305 (2014)
25. H. Benhar, A. Idrí, J.L. Fernández-alemán, Computer Methods and Programs in Biomedicine Data preprocessing for heart disease classification: a systematic literature review. *Comput. Methods Programs Biomed.* **195**, 105635 (2020)
26. G. Grekousis, Artificial neural networks and deep learning in urban geography: a systematic review and meta-analysis. *Comput. Environ. Urban Syst.* **74**, 244–256 (2019)
27. D. Elavarasan, D.R. Vincent, V. Sharma, A.Y. Zomaya, K. Srinivasan, Forecasting yield by integrating agrarian factors and machine learning models: a survey. *Comput. Electron. Agric.* **155**, 257–282 (2018)
28. J. Read, L. Martino, J. Hollmén, Multi-label methods for prediction with sequential data. *Pattern Recognit.* **63**, 45–55 (2017)
29. S. Shayaa et al., Sentiment analysis of big data: methods, applications, and open challenges. *IEEE Access* **6**, 37807–37827 (2018)
30. Y. Guo, L. Du, J. Chen, Max-margin multi-scale convolutional factor analysis model with application to image classification. *Expert Syst. Appl.* **133**, 21–33 (2019)
31. X. Li, H. Tang, D. Zhang, T. Liu, L. Mao, T. Chen, Histopathological image classification through discriminative feature learning and mutual information-based multi-channel joint sparse representation. *J. Vis. Commun. Image Represent.* **70**, 102799 (2020)
32. A. Mosavi, M. Salimi, S.F. Ardabili, T. Rabczuk, S. Shamshirband, A.R. VarkonyiKoczy, State of the art of machine learning models in energy systems, a systematic review. *Energies* **12**(7) (2019)
33. G. Sun, X. Zhang, X. Jia, J. Ren, A. Zhang, Y. Yao, Deep fusion of localized spectral features and multi-scale spatial features for effective classification of hyperspectral images. *Int. J. Appl. Earth Obs. Geoinf.* **91**, 102157 (2020)

# Review of Machine Learning Techniques for Analysis of Medical Data Sets



Deepali Bhende and Gopal Sakarkar

**Abstract** The biggest obstacles for the healthcare organizations are the lack of effective data governance procedure. The medical data should be clean, correctly formatted and precise to be used in the prediction process. There is an increasing gap between healthcare costs and outcomes. The reasons may include less usage of available evidences, poor management of insights from research, poor communication with patients which may lead to miss opportunities, resources wastage and probable harm to patients. To overcome this gap there is a need of continuous learning healthcare system. To achieve the improvement in the availability and quality of the data virtuous cycle should be formed between research and healthcare activities. Machine learning is solely creating healthcare smarter. It is playing an important role in the field of disease diagnosis. The main focus is on knowledge analysis. This knowledge is generated through the patient health data like X-ray results, vaccinations, blood samples, and many more. Machine learning models have the ability to extract the knowledge from the data which is very much useful in early prediction of the disease. For the treatment of any disease the identification of that disease is very much important task. In this analysis, there is a review of various conference and journal papers; that used Machine Learning techniques in disease diagnosis between 2018 and 2021. This research will definitely provide information about various machine learning techniques and their application in disease diagnosis.

**Keywords** Machine learning · Classification algorithm · Decision trees · KNN · SVM

---

*Present Address:*

D. Bhende (✉)

Department of Computer Science, G. H. Raisoni University, Saikheda, Chhindwara, India  
e-mail: [deepali.bhende.phdcs@ghru.edu.in](mailto:deepali.bhende.phdcs@ghru.edu.in)

G. Sakarkar

Department of Artificial Intelligence, G H Raisoni College of Engineering, Nagpur, India  
e-mail: [gopal.sakarkar@raisoni.net](mailto:gopal.sakarkar@raisoni.net)

## 1 Introduction

Machine learning (ML) is a branch of Artificial intelligence which focuses on number of methods that allow the investigator to learn and improve from experience [1]. These methods may permit analysis of large real-world databases. ML uses programmed algorithms that can access data and use it to learn from themselves. It uses historical data as input to predict new output values. ML algorithms are often categorised by how an algorithm learns to become more accurate in its predictions. There are four basic approaches: supervised, unsupervised, semi-supervised and reinforcement learning [2].

### 1.1 Supervised Machine Learning Algorithms

In this type of machine learning, the machine is trained by example. The data scientist supplies the machine learning algorithm with labelled training data. Both the input and output of the algorithm are specified [3]. The scientist makes corrections to the results generated by algorithm and this process is repeated over time till the algorithm obtained high level of accuracy or performance. There are three types of supervised learning as Classification, Regression and Forecasting.

**Classification:** The classification algorithm uses a technique that is used to identify the category of new observations on the basis of new training data. After learning the classification algorithm classifies new observations into number of classes and groups, such as Spam or not Spam, cat or dog, etc. [4].

**Regression:** It helps in finding the correlation between dependent and independent variables and enables to predict the continuous output variable based on the one or more independent variables. It predicts real or continuous values such as salary, temperature, age, price, etc. [4].

**Forecasting:** Forecasting uses historical data as input to make the informed estimates that are predictive in determining the direction of future trends [4].

### 1.2 Unsupervised Machine Learning Algorithms

For the identification of the different patterns the machine learning algorithm studies different features. There is no instructor to provide instruction. By analyzing available data machine determines the correlations and relationships between data. The ML algorithms interpret large data sets and perform the analysis. The structure of data is described by arranging it in some way. This is same as arranging or grouping the data into clusters [5]. As the dataset size increases, its capability to make decisions is also improves. This results in more refined output. There are two basic approaches:

**Clustering:** Clustering is one of the machines learning approach which groups the unlabelled dataset. Clustering is useful for segmenting data into number of groups and performing analysis to find different patterns on each dataset.

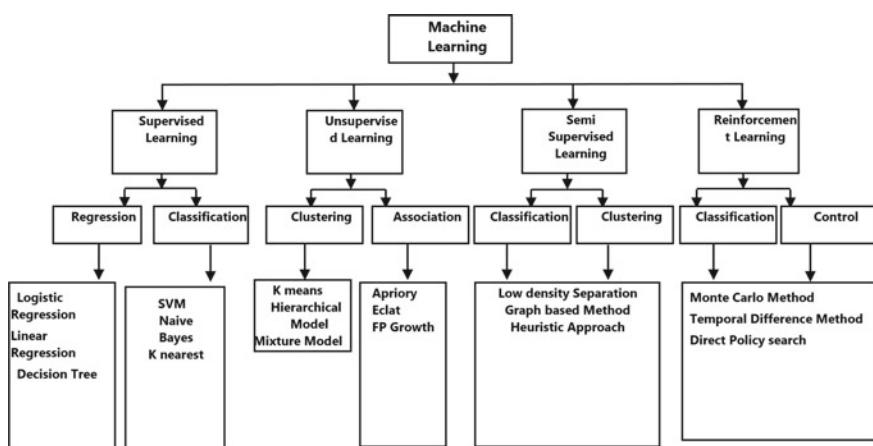
**Dimension reduction:** Mostly reduction techniques are used for reducing the number of variables in training data. Dimension reduction is followed by representation of data to extract the meaningful properties of original data [6].

### 1.3 Semi Supervised Machine Learning Algorithms

In Semi-supervised ML algorithm both labeled and unlabelled data can be used. Meaningful tags are assigned to Labeled data so that the algorithm can understand the data; on the other side the data without label does not consist of that information. This combination is used by ML techniques to assign the label to unlabeled data [7].

### 1.4 Reinforcement Learning Algorithms

This learning uses regimented learning processes. In this technique a set of actions, parameters and end values are provided to the machine learning algorithm. The machine learning technique tries to achieve optimal results by monitoring and evaluating each result. It first defines the rule and then explores all the possibilities. The technique uses trial and error method. The best results are obtained by using past experience as it is used to adapt with situation (Fig. 1).



**Fig. 1** Hierarchy of machine learning models [8]

## 2 Literature Review

Espino et al. [9] in 2021 have presented a model for the classification and prediction of diabetes. For the classification purpose different ML algorithms are used as logistic regression, random forest and multilayer perceptron. The study was extended with some more techniques as linear regression (LR), long short-term memory (LSTM) and moving averages (MA) method. PIMA Indian Diabetes Dataset is used for the analysis purpose. Finally the study concluded that MLP achieved an accuracy of 86.083% and LSTM achieved an accuracy of 87.26%.

Machine learning algorithms have been used by Jamkhandikar and Priya [10] to model a thyroid prediction and diagnosis. For Multicast thyroid dataset classification different algorithms as Naive Bayes, support Vector Machine and k-nearest neighbor were used. The proposed system shows that KNN (accuracy = 85%) gives better performance as compared to other algorithms.

The prediction model has been explained by Neelaveni and Devasana [11]. The model was based on the machine learning algorithm and used to detect the disorder of Alzheimer's disease. Different features such as MMSE, age, and education were considered. The model was used two approaches Support Vector Machine and Decision Tree. The comparison results shows that SVM gives better performance (85% accuracy) as compared to DT whose accuracy is 83%.

In 2020, Oyewo and Boyinbode [12] have presented an ensemble model for the detection of prostate cancer in Men. The ensemble model was based on Support Vector Machine, Decision Tree and Multiplayer Perceptron techniques. In the study the ensemble model was compared with RF, LR and ANN. Among all the individual classifiers the ensemble model gives best results.

Banerjee and Das [13] have presented a model which is based on digital image processing and machine learning algorithms for identification if lung cancer in its early stage. After extracting the attributes from the images different ML techniques such as ANN, SVM, RF were trained. The classification result categorized the tumor as benign or malignant.

Researchers in [14] have developed a model for prediction of Cardio Vascular Disease (CVD). The model depends on the machine learning ML classification algorithms like the logistic regression, decision tree, random forest, KNN and SVM. As compared to other classifiers Random Forest algorithm has given highest accuracy.

Emu et al. [15] was enhanced a technique for diagnosis of the level of liver fibrosis using Machine learning algorithms. Predictions of stages and generation of liver fibrosis rules was done with the help of important. In the study different approaches such as MLP, RF, and LG are evaluated and compared.

Vidya and Karki [16] in 2020 were developed a model which uses different ML algorithms as KNN, Naive Bayes and SVM. In this approach the skin lesion images were down loaded from ISIC. Preprocessing is applied on the images and feature extraction is done. The SVM ML technique gives the best results as compared to other techniques.

In 2019 Researchers in [17] were presented a model to detect heart disease. The analysis of heart syndrome is done with several parameters such as heart rate, age, diabetes, blood pressure and gender. Among different algorithms as SVM, DT, Random Forest, KNN and Logistic Regression, best results have been generated by KNN algorithm.

Selvathi and Suganya [18] in 2019 were proposed a model based on the Machine learning techniques. The data set with 283 infrared thermal images of eye has been selected as an input. Statistical features have been extracted from RGB images and texture features has been extracted from Gray images using Gray level co occurrence matrix. These features have been given as input to SVM machine learning algorithm for the detection of diabetic diseased eye.

Different image characteristics of brain such as gliomatosis, multifocal and multi centric tumors were studied by Cinarer and Emiroglu [19]. Using different statistical characteristics classification has been performed by using different ML approaches as SVM, KNN, random forest and linear discriminant analysis. Among all these SVM has given best results.

A diagnosis model for diagnosing hepatitis disease was presented by Islam [20]. The model was implemented using Waikato Environment for Knowledge Analysis software. Different ML techniques, such as KNN, Naive Bayes and RM (Random Forest) were used in the method. By using tenfold cross validations accuracy obtained was 98.6%, 93.2% and 95.8% for random forest, naive Bayes and K-nearest, respectively.

The Liver disease prediction model was built by Thirunavukkarasu and Singh [21] in 2018. In this work different classification techniques such as KNN, SVM and Logistic regression are used. To compare the performance different performance metrics were used as confusion matrix and accuracy score. From the result it is concluded that Logistic regression is best for the prediction of Liver disease.

The researchers in [22] were presented a model to diagnose breast cancer. For the diagnosis purpose they use an adaptive ensemble voting. The aim of the study is to detect the breast cancer. In the study artificial neural networks (ANN) and logistic algorithms performance has been compared with the constraint of minimum features. For this study the Wisconsin Diagnosis Breast Cancer dataset is used. ANN method with a logistic algorithm achieved 98.50% accuracy as compared to literature related work.

The Machine Learning approach was used by Tyagi et al. [23] for the diagnosis of thyroid disease. In the disease detection process ML plays an essential role. This work deals analysis of thyroid data depending on the data obtained from UCI repository. The diagnosis work consider different attributes as age, sex, TSH, T3, TBG, T4U, TT4 and FTI. Different machine learning techniques were used to identify risk of thyroid disease as ANN, KNN, SVM and DT. Among all these algorithms SVM gives best performance. As a conclusion the work suggested decreasing the number of parameters so that patient has to undergo less number of clinical tests which is both cost effective and time consuming (Table 1; Fig. 2).

**Table 1** Summary of different machine learning techniques

References	Year	Disease	ML algorithms	Dataset	Accuracy	Finding
[9]	2021	Diabetes	LSTM, MLP	PIMA Indian diabetes dataset	MLP = 86.083% LSTM = 87.26%	As compared to logistic regression and random forest MLP and LSTM give better results
[10]	2020	Thyroid disease	KNN, Naive Bayes, SVM	UCI repository	KNN = 85% SVM = 82% Naive Bayes = 83%	The proposed system shows that KNN gives better performance as compared to other algorithms
[11]	2020	Alzheimer	DT, SVM	Alzheimer's disease neuroimaging initiative (ADNI) database	SVM = 85% Decision tree = 83%	The proposed system shows that SVM gives better performance as compared to decision tree
[12]	2020	Prostate cancer	DT, SVM, MLP	github.com	DT = 92% SVM = 67% MLP = 97% EnsModel = 99.06%	The ensemble model developed using voting system gives highest performance
[13]	2020	Lung cancer	ANN, SVM, RF	UCI repository	SVM = 80% RF = 70%, ANN = 96%	Among the different three models ANN gives highest results

(continued)

**Table 1** (continued)

References	Year	Disease	ML algorithms	Dataset	Accuracy	Finding
[14]	2020	Cardio vascular	DT, SVM, RF, LR, KNN	NIDDK (National Institute of Diabetes and Digestive and Kidney)	SVM = 77.14 LR = 74.28% DT = 74.28% KNN = 68.57% RF = 85.71%	In the prediction of cardiovascular disease among all other classifiers Random Forest gives the better results
[15]	2020	Liver fibrosis	RF, MLP, LR	University Ain Shams, At ElDemerdash Hospital	LR = 97% RF = 97.228% MLP = 98%	The study shows that MLP classifier obtained the highest accuracy rate
[16]	2020	Skin cancer	KNN, SVM	1000 instances were collected from ISIC 2017	KNN = 86.25% SVM = 97.8%	The proposed system shows that SVM classifier outperforms with 97.8% accuracy which is better than KNN
[17]	2019	Heart disease	KNN, SVM, RF, Naïve Bayes	Kaggle dataset	SVM = 82% Naïve Bayes = 80% K-NN = 85% RF = 81% DT = 82%	The proposed system shows that KNN is more effective algorithm as compare to other algorithms

(continued)

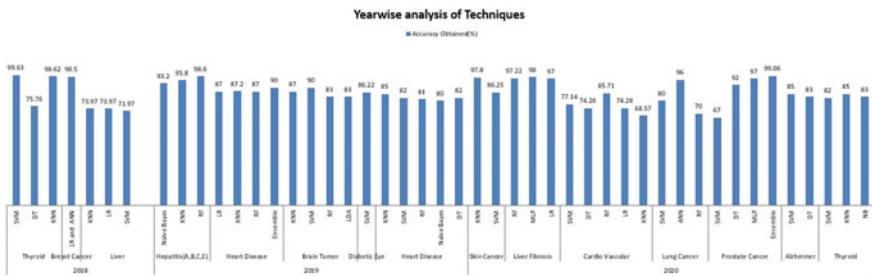
**Table 1** (continued)

References	Year	Disease	ML algorithms	Dataset	Accuracy	Finding
[18]	2019	Diabetic eye	SVM	FLIR.com 283 eye thermal images	SVM = 86.22%	The study represented the use of SVM techniques to classify the thermal images as normal or diabetic diseased eye
[19]	2019	Brain tumors	KNN, SVM, RF, LDA	Total patients 33 and 20 images (on an average) per patient	SVM = 90% RF = 83% KNN = 87% LDA = 83%	Among all other classifiers SVM gives highest performance
[20]	2019	Hepatitis-A, B, C, E	Naïve Bayes KNN, RF	Real data of hepatitis patients	Naïve Bayes = 93.22% KNN = 95.8% RM = 98.6%	Random forest classifier gives better results for the diagnosis of hepatitis virus
[21]	2018	Liver	KNN, LR, SVM	Indian liver patient dataset (ILPD)	KNN = 72.97% LR = 73.97% SVM = 71.97%	Due to highest sensitivity LR is more suitable Model as compared to KNN
[22]	2018	Breast cancer	Ens model using ANN and LR	Wisconsin Breast Cancer Database	EnsModel = 98.5%	Combination of two classifiers gives better results as compared to individual technique

(continued)

**Table 1** (continued)

References	Year	Disease	ML algorithms	Dataset	Accuracy	Finding
[23]	2018	Thyroid	DT, SVM, KNN	UCI repository	DT = 75.76% KNN = 98.62% SVM = 99.63%	The proposed system shows that SVM algorithm gives best performance as compared to other algorithms



**Fig. 2** Year wise analysis of various machine learning techniques

### 3 Review Findings

In the current literature review it is observed that the most widely used Machine Learning algorithms are Support-vector machines, Logistic regression, Linear discriminant analysis, K-nearest neighbor algorithm, Naive Bayes, Decision Trees and Neural Network. We need some way to measure the quality of fit i.e. how well its predictions actually match the observed data. Secondly most of the researchers conducted a research on UCI Dataset. In some papers KEEL, Kaggle, GitHub and Figshare are also mentioned. In order to exploit all the potential and usefulness of the datasets more attention is required to pay for disease classification and prediction. There is need to check possibility to design a new dataset with proper attributes from the real world. The scope is also available to check the fact that whether the addition of more instances to dataset leading to more strong results or not.

### 4 Conclusion

Overall medical expenditure can be reduced using the Machine learning approach in medical diagnosis. The results obtained by the algorithms can be helpful as a second opinion for doctors. For making decision mathematical models are used. In medical diagnosis it is very much important to improve the accuracy with which a disease is detected the primary focus of this review is to stresses the use of ML for effective medical diagnosis. It is observed that over the years the use of ML in medical diagnosis has increased gradually. For effective analysis and prediction of different chronic diseases, extraction of knowledge from huge patient's dataset is required and it becomes easy by applying Machine learning algorithm. So in this regard researchers are helpful to the needy persons, doctors and society.

## References

1. N. Caballe-Cervigon, J.L. Castillo-Sequera, J.A. Gomez-Pulido, J.M. Gomez-Pulido, M.L. Polo-Luque, Machine learning applied to diagnosis of human diseases: a systematic review. *Appl. Sci.* **10**(15), 5135 (2020)
2. G. Battineni, G.G. Sagaro, N. Chinatalapudi, F. Amenta, Applications of machine learning predictive models in the chronic disease diagnosis. *J. Pers. Med.* **10**(2), 21 (2020)
3. A.M.U.D. Khanday, S.T. Rabani, Q.R. Khan et al., Machine learning based approaches for detecting COVID-19 using clinical text data. *Int. J. Inf. Technol.* **12**, 731–739 (2020)
4. M. Ayad, H. Kanaan, M. Ayache, Diabetes disease prediction using artificial intelligence, in *21st International Arab Conference on Information Technology (ACIT)*, Giza, Egypt, 2020, pp. 1–6
5. I.V. Derevitskii, S.V. Kovalchuk, Machine learning-based predictive modeling of complications of chronic diabetes. *Procedia Comput. Sci.* **78**, 74–283 (2020). ISSN 1877-0509
6. T. Egerton, E. Sochima, F. Deedam-Okuchaba, A model to detect heart disease using machine learning algorithm. *Int. J. Comput. Sci. Eng.* **7**, 1–5 (2019). <https://doi.org/10.26438/ijcse/v7i11.15>
7. N.M.H. Tayarani, Applications of artificial intelligence in battling against covid-19: a literature review. *Chaos Solitons Fractals* **142**, 110338 (2021)
8. A. Ray, A.K. Chaudhuri, Smart healthcare disease diagnosis and patient management: innovation, improvement and skill development. *Mach. Learn. Appl.* **3** (2021). ISSN 2666-8270
9. D. Espino, U.M. Butt, S. Letchmunan, M. Ali et al., Machine learning based diabetes classification and prediction for healthcare applications. *J. Healthc. Eng.* **2021**, Article ID 9930985 (2021)
10. D. Jamkhandikar, N. Priya, Thyroid disease prediction using feature selection and machine learning classifiers. *Int. J. Anal. Exp. Modal Anal.* **XII(VII)** (2020). ISSN: 0886-9367
11. J. Neelaveni, G. Devasana, Alzheimer disease prediction using machine learning algorithm, in *6th International Conference on Advanced Computing & Communication Systems (I.C.A.C.C.S.)* (2020)
12. O.A Oyewo, O.K. Boyinbode, Prediction of prostate cancer using ensemble of machine learning techniques. *IJACSA—Int. J. Adv. Comput. Sci. Appl.* **11**(3) (2020)
13. N. Banerjee, S. Das, *Prediction Lung Cancer—In Machine Learning Perspective* (IEEE, 2020)
14. N.K. Kumar, G.S. Sindhu, D.K. Prashanthi, A.S. Sulthana, Analysis and prediction of cardiovascular disease using machine learning classifiers, in *6th International Conference on Advanced Computing & Communication Systems (I.C.A.C.C.S.)* (2020)
15. M. Emu, F.B. Kamal, S. Choudhury, T.E. Alves de Oliveira, *Assisting the Non-invasive Diagnosis of Liver Fibrosis Stages Using Machine Learning Methods* (IEEE, 2020)
16. M. Vidya, M.V. Karki, *Skin Cancer Detection Using Machine Learning Techniques* (IEEE, 2020)
17. N. Basha, A. Kumar, G. Krishna, P. Venkatesh, Early detection of heart syndrome using machine learning technique, in *4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT)* (2019)
18. D. Selvathi, K. Suganya, *Support Vector Machine Based Method for Automatic Detection of Diabetic Eye Disease using Thermal Images* (IEEE, 2019)
19. G.G. Cinarer, B.G. Emiroglu, Classification of brain tumors by machine learning algorithms, in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (2019), pp. 1–4. <https://doi.org/10.1109/ISMSIT.2019.8932878>
20. T.I. Trishna, S.U. Emon, R.R. Ema, G.I.H. Sajal, S. Kundu, T. Islam, Detection of hepatitis (A, B, C and E) viruses based on random forest, K-nearest and Naïve Bayes classifier, in *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (2019), pp. 1–7. <https://doi.org/10.1109/ICCCNT45670.2019.8944455>
21. Thirunavukkarasu, A.S. Singh, M. Irfan, A. Chowdhury, Prediction of liver disease using classification algorithms, in *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, 2018, pp. 1–3. <https://doi.org/10.1109/ICCCA.2018.8777655>

22. N. Khuriwal, N. Mishra, Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm, in *IEEMA Engineer Infinite Conference (eTechNxT)* (2018), pp. 1–5. <https://doi.org/10.1109/ETECHNXT.2018.8385355>
23. A. Tyagi, R. Mehra, A. Saxena, Interactive thyroid disease prediction system using machine learning technique, in *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (2018), pp. 689–693. <https://doi.org/10.1109/PDGC.2018.8745910>

# Incorporating Contextual Information in Prediction Based Word Embedding Models



K. Vimal Kumar, Shruti Ahuja, Skandha Choudary, and Pritee Parwekar

**Abstract** A clustering-based language model is proposed for analyzing the performance of context sensitive word embedding models that uses contextual information. The construction of text readability and prediction models faces several shortcomings due to the complex nature and structure of language. The language structure is complicated all the more by words having vastly different meanings and interpretation based on the context in which it is used. This paper aims to resolve this issue by first clustering the sentences based on similarity and then performing word embedding separately on each of these clusters to obtain an enhanced outcome. This would serve to embed the same word separately in varying context as an improvement over the standard existing word embedding models provided by various prediction based models. Comparing the two approaches, our results have showed that clustering improves the performance of the model and discriminate the contextual information based on sense which leads to more accurate representation in vector form.

**Keywords** Word embedding · Sense based embedding · Word2Vec · K means clustering · Birch clustering

---

K. Vimal Kumar (✉)  
Jaypee Institute of Information Technology, Noida, U.P., India  
e-mail: [vimalkumar.k@gmail.com](mailto:vimalkumar.k@gmail.com)

S. Ahuja  
Amazon, Hyderabad, India

S. Choudary  
AXA-XL, Gurgoan, India

P. Parwekar  
SRM Institute of Science & Technology, Ghaziabad, U.P., India

## 1 Introduction

Natural language processing has innumerable applications in various industries. Companies use it to process large amounts of data and documents for the identification and analysis of useful data. However, it is not possible for a computer to naturally understand human language with its complexity, ambiguity as well as its variations in different contexts. To make the task of understanding the language similar, it is imperative to process the linguistic data and somehow structure it meaningfully for better understanding.

Word embedding is the technique of processing text and mapping, or rather, embedding words with vectors that can be comprehended by the system for further processing in applications like machine translations, sentiment analysis, text prediction, etc. The model allows the creation of vectors for each individual word, where similar words are mapped closer to each other so that one can figure out the correlation between words like ‘girl’ and ‘boy’ or numbers like ‘one’ and ‘two’.

The basic model used for word embedding is the prediction based language model which uses neural networks to group words inside a vector space. Two types of word embedding models are used for the purpose: Continuous Bag of Words and Skip Gram Model. While these models have proven to be successful in interpreting the text and creating vectors for semantically similar words and prediction, they leave much to be desired in unraveling the ambiguousness of language where the same word can have different meanings and interpretations depending on its use. For e.g., the word ‘apple’ can be interpreted as both, a fruit, as well as the name of the multinational technology giant that sells consumer electronics. Similarly, the word ‘bank’ when used in a sentence that alludes to a river often refers to a river bank, but when used in a sentence that includes money and finance, can mean the traditional banks where financial transactions are carried out.

The aim of this proposed work is to make a word embedding model that is more accurate as well as adaptive to the quirks of language and ready for use in real-world applications. For that purpose, a new approach is suggested which includes clustering the sentences based on the context used. The clustering techniques used are K-Means and Birch Clustering. Once different clusters are formed, the prediction based language models are applied on individual clusters, hence creating a model where a single word can have different vectors based on the subject usage.

This approach improves its accuracy and increases the probability of semantically mapping closer words next to each other in multiple case.

## 2 Related Work

Sun et al. [1] achieved polysemous word embedding by assigning a global base vector to each word and then modeling the contextual vector based on its linear combination of contextual weights that were calculated from the co-occurrence of

words in specific situations. The approach was then evaluated on several parameters to find the relevance of the word embedding using Spearman's Correlation. The authors successfully developed an approach that can generate vectors of any word for any context as well as compete with other established unsupervised methods.

Cha et al. [2] worked on developing a model that can effectively discover semantic similarity in sentences. The model works by performing clustering on word embedding to create a language model that can predict text and has been evaluated to be far superior to the general bag of words models and the like using the Common Core Corpus and the Wiki-SimpleWiki corpus.

Yang et al. [3] accomplished in creating a multi-sense model for neural machine translation by incorporating a different sense specific embedding for each sense type of a word. It uses an approach based on a latency controlled bidirectional RNN model for the sense search module to compute semantic and syntactic information.

Mikolov et al. [4] proposed the Continuous Bag of Words and Skip Gram models which have been used to create context-based words based on the words closest to each other. Continuous Bag of Words uses multiple contexts to predict a single word whereas the Skip Gram Model makes predictions based on a single word given as an input to a linear classifier layer and the surrounding words are predicted and projected as an output. Their previous work [5] also compared the two proposed models and calculating the most efficient text prediction model for large datasets for use in industrial applications and the development of further advanced models.

Zhang et al. [6] hypothesized BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) approach for clustering of large data sets. It works as an improvement over the traditional hierarchical clustering method and has been used for sensebased sentence clustering in this paper.

### 3 Implementation

A two-step approach has been developed for sense based word embedding in this paper. The steps are given as follows:

1. Firstly, the raw text is taken and clustered sentence-wise, using two different models on the same text, the first being K Means Clustering and the second, Birch Clustering model. Each model creates 8 different clusters of the text.
2. On each of the generated clusters, the two word2vec models—Continuous Bag of Words (CBOW) and Skip Gram have been separately applied and the loss for each of them has been calculated and detailed in Sect. 4.

#### 3.1 Clustering

For clustering to work on a text data set, a python dictionary is used for the mapping of words to create a sparse matrix where the frequency of occurrence of words is

used for assignment of values. The values are later computed again using Inverse Document Frequency vector. Here, the weighted average of the count word vectors is taken for each sentence and this sentence average is projected to a Euclidean semantic space, providing us data in a form on which we can perform clustering.

**K Means Clustering:** K means clustering is an unsupervised learning mechanism to group unprocessed data into clusters based on similarity or closeness. The term K defines the number of clusters that are used in the algorithm and it is a user-defined value. The main feature of K Means Clustering is a centroid which is a collection or rather, the average of the values of all the points in a cluster and can be used for labeling of new data. It works by recursively computing the centroids of a cluster and assigning the cluster with the nearest centroid to the grouped data until a stable value is reached. The algorithm uses a Euclidean function to compute distance.

Number of points  $x_1, x_2, x_3, \dots, x_n$  are assigned to clusters  $c_1, c_2, c_3, \dots, c_k$  according to the formula:

$$c_i(x_n) = \min ||c_{ni} - x_n||^2$$

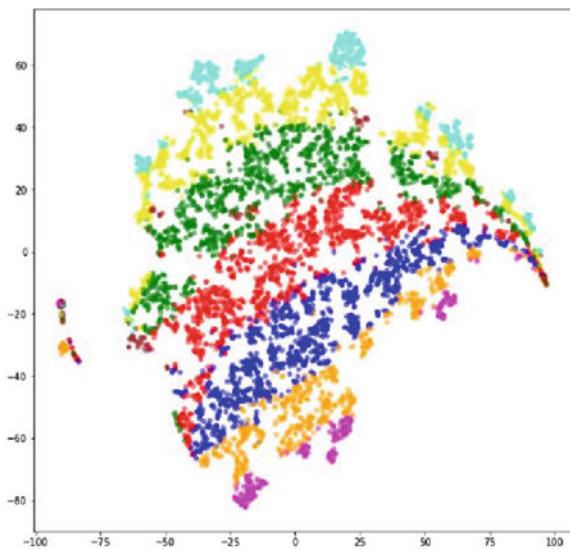
where,  $c_{ni}$  is the centroid of cluster  $c_i$ .

**Birch Clustering:** It is an unsupervised learning algorithm used to process extremely large data and apply hierarchical clustering. The greatest benefit of the algorithm is its ability to dynamically cluster the sentences provided to it, making it both memory and time efficient as compared to K Means, which requires multiple reiterations for final computations of the centroid. The basic algorithm works by building a CF tree of  $n$  points in data by computing the average and the square sum average of the values of the data points [6]. With two parameters, branching factor and threshold value defining the tree, the algorithm automatically presents the clusters and sub-clusters as branches of this tree (Figs. 1 and 2).

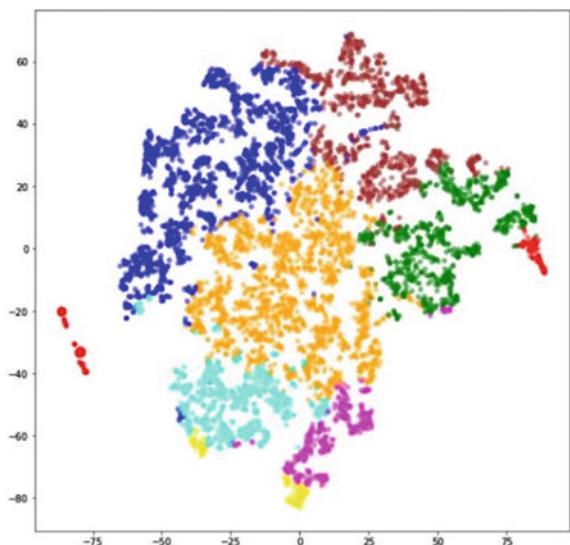
### 3.2 Word Vectorization Using Word2vec

The use of vector space models to represent vectors in a semantic space and mapping words similar words close to each other is a distinguishing feature of word embedding. The basic hypothesis governing our function is that words appearing in the same context are similar, which when interpreted in reverse can mean that words found in different situations are semantically different. However, when the same words are discovered to form a pattern by appearing in two contrasting contexts with no semantic similarity, we can conclude that the same word can have distinct meanings or sense. Once, the sentences are clustered according to their semantic similarity, word2vec models can be applied on each of these clusters. Both models make use of the word count and co-occurrence with other words to compute the value in vector space.

**Fig. 1** Clusters formed by K means clustering algorithm



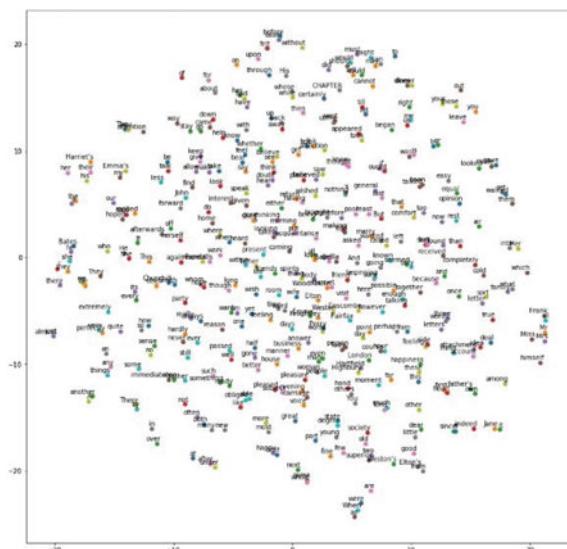
**Fig. 2** Clusters formed by Birch clustering algorithm



**Continuous Bag of Words (CBOW):** The main feature of this model is its ability to predict a single word given a set of multiple words as context and taking their average gradient. This model evaluates the set of neighboring words as its context and gives us a log likelihood of the occurrence of another word when the input is given. It uses shallow neural networks to create a hidden layer to compute the weight of words in the weight matrix.

**Skip-Gram Model:** It has a different approach when compared to the CBOW model as it uses the normal distribution function to predict our target word or set of target words when given a single context word. Its advantage lies in its ability to capture two different senses for the same word and giving two different vector representations for the same word. A skip window is defined which trains the model to correctly predict the target words and optimize the function successively (Figs. 3 and 4).

**Fig. 3** Continuous bag of words mapping in vector space



**Fig. 4** Skip Gram model mapping in vector space



## 4 Result Analysis

We have used a standard corpus from gutenberg.org to train our model and perform analysis. Firstly, the raw text has been preprocessed by being tokenized in the form of sentences and then stripped of all punctuation. Later, word tokens were formed from the sentence vectors so that the words can be encoded and projected into vector space while retaining the basic features of a sentence. After the words were vectorized into a Euclidean semantic space, an average of the word vectors was taken for each space, and the sentence coordinates were calculated. This was the information that was used for the clustering of sentences under different labels. The two different models that were used were K Means Clustering and Birch Clustering.

**Silhouette Score:** The value of this score measures from  $-1$  to  $+1$ . Silhouette score measures the similarity of an object to the cluster it is assigned to. The lower the silhouette score, the lower the similarity and lower the distinction between two different clusters. Silhouette Score can also be measured in terms of a function  $S(i)$ . Here,

$$S(i) = \text{Avg}(i) - \text{Min}(i) / \max(\text{Avg}(i), \text{Min}(i))$$

where,

$S(i)$  is the silhouette score of an object  $i$  in a cluster,

$\text{Avg}(i)$  is the average distance of object  $i$  from all the other objects in the cluster,

$\text{Min}(i)$  is the smallest distance of object  $i$  from all points in the clusters which object  $i$  is not a member of.

For each of these models, the silhouette score shown in Table 1 is an average of the silhouette score calculated for all the objects in the clusters.

After the two different clustering algorithms are applied on the dataset, the data is divided into 8 different clusters. At this point, word2vec models are applied on each of these clusters. The accuracy of the models has been determined using a loss function that calculates the loss sustained in each successive operation that optimizes the model. The loss in accuracy of the model gives a good indicator of the performance of the model. Previously, it had been determined that Continuous Bag of Words model performs more efficiently than the Skip Gram Model. The same result has been illustrated in the findings of our research (Table 2).

Tables 3 and 4 clearly show that the losses of the data are marginally reduced after clustering them based on the context of the sentences. This proves the hypothesis of

**Table 1** Silhouette Score showing the accuracy of clustering algorithms used

Clustering algorithm	Silhouette score
K means	0.287
Birch	0.229

**Table 2** Loss function of the word2vec models for the entire corpus

Word2vec model	Average loss
Skip Gram model	2.29
CBOW model	1.03

**Table 3** Loss function of the word2vec models for the clusters formed using K means

	Average loss for K means	
	CBOW	Skip Gram
Cluster 1	0.58	1.68
Cluster 2	0.52	1.84
Cluster 3	0.33	1.65
Cluster 4	0.57	1.92
Cluster 5	0.58	1.97
Cluster 6	0.10	1.50
Cluster 7	0.49	1.90
Cluster 8	0.22	1.55

**Table 4** Loss function of the word2vec models for the clusters formed using Birch clustering

	Average loss for Birch	
	CBOW	Skip Gram
Cluster 1	0.70	2.11
Cluster 2	0.01	0.37
Cluster 3	0.04	0.77
Cluster 4	0.40	2.11
Cluster 5	0.42	1.79
Cluster 6	0.75	2.17
Cluster 7	0.31	1.65
Cluster 8	0.40	1.81

sense based word embedding and allows an improvement over the traditional word embedding models to correctly target words according to their sense.

## 5 Conclusion

The goal of this research was to develop a model that correctly embeds words into vectors based on their sense. By clustering the sentences according to their context and then applying prediction based models separately on the clusters, we were able to create a new vector for each sense type of a word. This can be used in any natural language applications where there is need for context sensitive information. Since

this model is based on contextual information established through clustering, it can be implemented and extended for any other languages too.

## References

1. Y. Sun, N. Rao, W. Ding, A Simple approach to learn polysemous word embeddings (2017)
2. M. Cha, Y. Gwon, H.T. Kung, Language modeling by clustering with word embeddings for text readability assessment (2017)
3. Z. Yang, W. Chen, F. Wang, B. Xu, Multi-sense based neural machine translation, in *2017 International Joint Conference on Neural Networks (IJCNN)*
4. Mikolov et al., Efficient Estimation of Word Representations in Vector Space (2013)
5. Mikolov et al., Distributed Representations of Words and Phrases and their Compositionality (2013)
6. T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: an efficient data clustering method for very large databases (1996)
7. J.E. Alvarez, A review of word embedding and document similarity algorithms applied to academic text (2017)
8. P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriching word vectors with subword information. *Trans. Assoc. Comput. Linguist.* **5**, 135–146 (2017)
9. W. Zhu, X. Jin, J. Ni, B. Wei, Z. Lu, Improve word embedding using both writing and pronunciation. *PLoS ONE* **13**(12), e0208785 (2018). <https://doi.org/10.1371/journal.pone.0208785>
10. Z.-L. Ye, H.-X. Zhao, Syntactic word embedding based on dependency syntax and polysemous analysis. *Front. Inf. Technol. Electron. Eng.* **19**, 524–535 (2018). <https://doi.org/10.1631/FITEE.1601846>

# Distributed Cloud Data Center for Efficient Services in Logistics Management



Shivani Dubey, Vikas Singhal, Rajanish Jain, and Amit Agrawal

**Abstract** An issue in research is a challenging area to be improved, eliminated and the troubling question which exist in theory, practice which needs to have meaningful investigation and information distribution. Logistics information system (LIS) is needed to drive globally decrease or increase by the nature. Thus it is required, LIS should be in the form of distributed system rather than centralized system. As there is availability of anywhere and anytime the cloud provider are providing resources for sharing information by using its services in logistics operations. Logistics under the umbrella of supply chain itself is distributed over various regions so it should also adopt distributed cloud datacenter for its efficient services. Thus the basic research problem under study is to provide an efficient and scalable solution via distributed datacenter over cloud for managing logistics.

**Keywords** Logistics management · Cloud computing · Load balancing algorithm · DSBP

## 1 Introduction

Logistics management has been very crucial factor and focused from researchers and developers in past few years. The application of inefficient and ineffective Logistics.

Information System (LIS) adversely affects the firm's Information capabilities which in turn affects the firm's financial performance as well [1]. In this context a number of uncertainty management issues arise like customer complaints, order cancellation, product damage in inventory and packaging, price variation in purchasing and procurement etc. which are required to be addressed adequately with

---

S. Dubey (✉) · V. Singhal · A. Agrawal  
Greater Noida Institute of Technology, Greater Noida, India  
e-mail: [dubey.shivani@gmail.com](mailto:dubey.shivani@gmail.com)

R. Jain  
Institute of Management Studies, Ghaziabad, India

the help of LIS [2–4]. The ability to be sustained is a very critical task in logistics management [5]. The operational performance of logistics has positive effect on performance of any company [6]. Logistics Management consider the several activities from services of customers, order generating, transportation, inventory process, warehousing, demanding, procurement, production, forecasting and distribution that all are helped by massive information flow [7]. The issues in information flow, participants and operation in company needs extensive research for operation implementation and cost minimization for products distribution in accurate way. The responsibility of cloud computing improves the success of LIS. A logistics network is a sequence of value added operations deal by different enterprises. The logistics information starts from the raw material supplier to users at the end level. Growth of product variation creates demand forecasting which is more difficult in multiple numbers of suppliers to manage the higher coordination cost. In this overall research, we give an overview of adoption of cloud based application for LIS to apply LIS over Cloud based data center by using load balancing techniques to ensure the better performance and gives better effects for users satisfaction. Cloud computing generally manages storage and computing resources in centralized system. In logistics management, multiple numbers of distribution centers are scattered in various regions. Centralizes Datacenter is being used by the centralized system for the track of delivering the information and services. Then network congestion will be created and this issue will be depended upon data center where several requests are occurred at the same time from several regions. At that time, load balancing techniques are required for reducing and improving the issues of overloading over the centralized data center. There are so many chances to increase the higher latency due to heavy load of requests for any particular service. Logistics is distribution process in any organization which wants to implement its supply process to require IT services at the higher level of supply chain. Distributed data center provides better environment to improve bandwidth and traffic for implementing supply chain over cloud. A logistics system is facing several issues in information distribution among the logistics partners, suppliers and users by using cloud services. Supplier, distributor and users need information sharing in accurate form which should be in real time. Since services of cloud computing provide highly efficient information sharing in real time. Hence it can be analyzed the requirement of cloud based applications for managing logistics over cloud [8]. In current scenario, suppliers, logistics partners and users expect fast processing of information at each level of logistics operation and minimize the response time without having any fault tolerance for downtime. Online slow processing of any application presents the result of dissatisfaction of users and has impact on company's business for establishes the relations with users and other businesses. High latency is to be analyzed and measured in milliseconds which have speed matter for connecting the users to the businesses. Higher latency has various issues related to data center distribution, inefficient routing approach, limited servers and network traffic. Load balancing approach is so much required globally due to higher latency among the suppliers and users. The locations of user and information at each level are not accessed or processed on data center/server nearest to them without having efficient and effective load balancing technique. When a data

center is not be in efficient environment, the service provider and users can suffer several issues through large response delays [9].

## 2 Logistics Information System

LIS plays a very important role in achieving logistics competitiveness. Most organizations do not consider the good strategies in improvement of the required information systems of this century. There is partially lacking to understand about the causal links between the characteristics of cooperation and priorities of LIS. The LIS allows people to evaluate actual information and also planning functions of information. The information system gives the facility to use planning functions that are also supported by a forecasting function. LIS controls overall functionality of the system and manages all the operations of logistics such as coordination and planning. Co-ordination and planning helps the area of user where an operation of supply chain is considered at the highest level to plan distribution of the products, promotions and services. The purpose of logistics operation is to share information among the suppliers, logistics partners and users in product tracking process. Logistics operations are based on order management and procurement. Order management describes accuracy in request entry and customer ordering with accurate information to improve purchasing order and ensure the better performance of overall supplier compliance [10]. When this type of advance technique is used specific software solution into the organization, the information accuracy, location accuracy and time accuracy becomes priority for information distribution among the suppliers, logistics partners and users. Cloud computing provides fast and forward software for providing accurate information with higher quality at accurate time and location. LIS provides accurate flow of information among the segments such as warehouses, inventories and transportations for providing good services for users. LIS is a tracking based process to identify the accurate information at all the levels of logistics at lower cost and lower time. Logistics have a significance responsibility for sharing information at each level of the organization. The major problem is to plan accurate approach for LIS to reduce the cost of information distribution. Mostly organizations are using LIS to capture the accurate information availability and in structured form so organizations can utilize the information whenever they need to distribute it at accurate time and location.

## 3 Cloud Computing

Cloud computing is a modern framework to allows to user to access on demand network access for sharing resources such as storage, networks, services, servers and applications) which can be released and reserved with minimum effort of management and interaction of service provider. Cloud computing always implements the

service delivery process at minimum cost as per requirement which made it advance for using the online virtual system, different platforms and storage effectively with minimum cost. Cloud computing includes different networks, software, service over web and resource distribution. There are different types of centralized, distributed servers, data centers and clients formed by cloud services. It involves minimum cost, minimum stress for user, high availability, flexibility, scalability and fault tolerance on demand based services. Cloud is a curtain over the internet in which internet can be hide and the cost of operations of cloud can be minimized over the internet [11]. Cloud computing is a modern technology which implements the process in which we create a new design for various applications, computing system and building of software for different services. It is a dynamic provisioning which designs different services for network storage, data centers and large computing capability. The resources over the internet are offered as per demand basis from different vendors of cloud services. Anyone can subscribe with online payment and also can access to any type of cloud service, configure server and deploy server for application time basis and access application as per demand basis for the specific time [12].

## 4 Research Findings

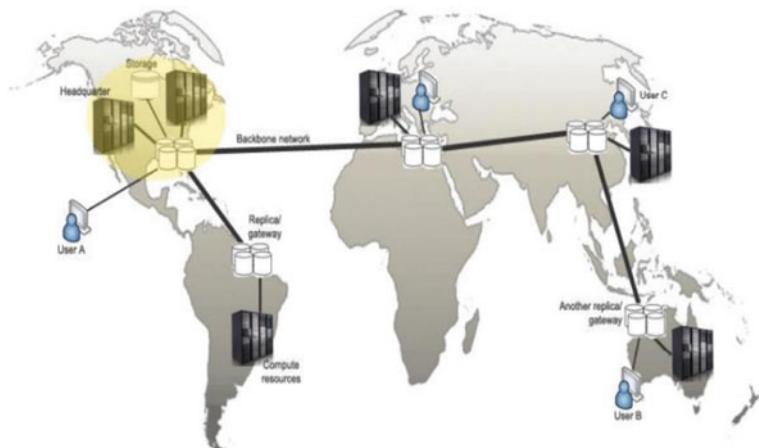
The objective of our research is to establish a new cloud based application for logistics coordination which is time efficient and cost effective solution logistics operations. In our research, we have compared three load balancing algorithms with our proposed algorithm and have implemented them for centralized and distributed data center. They are (1) Round Robin (2) Active Monitoring (3) Throttled Load Balancing and (4) DSBP Algorithm. Our research objective is to analyze the achievement of all these algorithms for logistics performance and improvement and to compare centralized data center (CDC) and distributed data center (DDC) that determine a strategy which gives cost effective and time efficient service for logistics information system. There are research objectives:

- (1) Information sharing between logistics partners and users efficiently
- (2) Selection of cloud service for logistics information system
- (3) Distributed Datacenter for logistics information system
- (4) Implementation of Service broker policy algorithms.

## 5 Issues in Centralized and Distributed Data Center

There are so many service for logistics industries provides by the cloud computing by using IaaS. IaaS allots the facilities to users for utilizing the resources for sharing information, executing information and high level computation. Generally data centers work as a servers including large scaling of information sharing, storage

based on remote server and executing the operations. The CDC provides the environment of centralization in which multiple data centers are connected to a single centralized data center for sharing information among the users in a scaling geographical area. CDC uses selection policy and proximity policy based on centralized approach over the cloud where a central admin gives the permission to VMs for utilizing the resource by cloud service provider. Sometimes CDC has multiple requests of users at higher range then the central admin tracks the shared information by using centralized policy. These requests increase probability of network traffic due to overload on the data center at the same time. The centralization approach is used for central admin having less working processing to maintain the CDC from inside or outside the region. The central admin also not has command on local manger that is connected to in some other branches for distributing the load on various VMs. This situation generates higher latency, large service demands and increases response time and data transfer cost. So in order to assume these issues in CDC, the novel load balancing approaches are required to implement the performance of logistics at higher level. DDC connects to the multiple data centers in various region in large geographical area. DDC can be risky for taking decision to adopt efficient distributed approach for information sharing among the users from different locations at the same time. This situation presents the poor impact of DDC for improving the latency, overload on the data center, response time of multiple user's request and data transfer cost over the cloud. We are taking an example of centralized data center of Facebook in Fig. 1, which is connected to the DDC at different locations. These DDC are very risky to find out the efficient solution for maintain the network traffic and bottleneck in network due to inefficient process of data centers. To resolve this issue Face-book presented content delivery network (CDN) for handling the problems related to DDC to maintain the bandwidth, storage, latency, memory, information sharing



**Fig. 1** CDN in Face book Distributed Data Center [13]

etc. in widely geographical area. CDN provided a system facility for accepting all the components those are still working in processing [13].

## 6 Design Methodologies

This research provides insights of the role of Distributed systems and Cloud Computing in Logistics. Logistics is described as segment of the supply chain process which considers planning, controlling, flow of information, storage, services and information from the source to destination for consuming requests of users. IT is described as an organizational enhancing ability to develop and maintain digitally-enabled interorganizational relationships. Cloud has higher capacities to use its applications and service for executing the requests of the user in systematic manner. Cloud gives an advance shape to information segment by using IT based services yet to come in year. LIS controls the processes of entire system and provides coordination to all its operations. Operations are used to share information among the suppliers and users for distributing products and information related to the product based on functionalities like ordering, implementing and controlling the procurement for suppliers and users.

With help of advance technologies and methods software can implemented in company for tracking information related to logistics in systematic manner. Cloud provides enhancement of IT based software for providing quality based applications for LIS. Load balancing technique solves the major problems in cloud to describe the ideal approach and method to minimize the overloading on each data center in cloud environment. Load balancing has distribution approaches to excess extra load over the data centers. The load balancing technique provides sufficient service for better utilization and provision to implement the performance of the data center. Load balancer has multiple incoming requests from multiple locations to distribute multiple data centers for maintaining accuracy in the distribution of load [14]. According to the various scheduling and load balancing algorithms, the major issues are investigated by proposed DSBP algorithm which has ability for better data transfer cost and response time. In significance this research considers the improvement of round robin, active monitoring and throttled load balancing and proposed my algorithm which leads to better resource utilization.

- (a) Round Robin Algorithm (RR): It divides processes among all the data centers. The process allocation system is managed locally individually from the remote data centers. It has fixed quantum time which creates time limitation. In this algorithm, the load balancer assigns VM for requesting data center in cycle manner among all connected data centers [15].
- (b) Active Monitoring Algorithm (AM): Active monitoring maintains the requests on every virtual machine and requests are currently generated to that virtual machine, when a request is located on new virtual machine. This algorithm revert backs virtual machine id to controller of data centers which sends the

request to the virtual machine identified by its id and also specifies newly allocated request and cloudlets has sent to it [16].

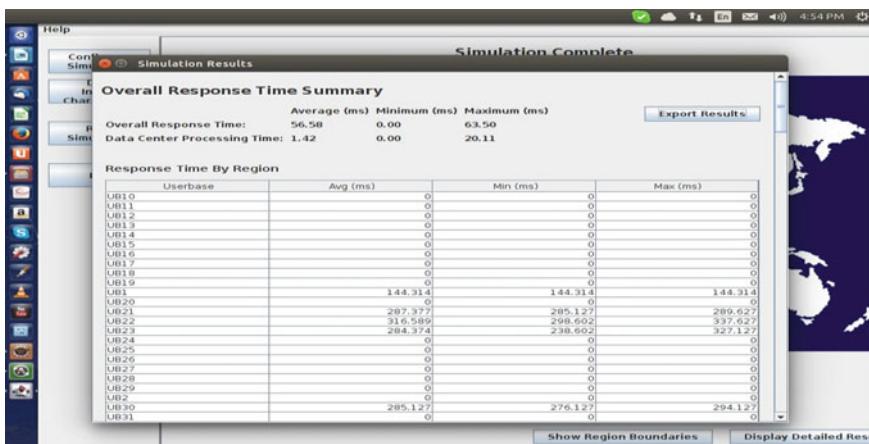
- (c) Throttled load balancing Algorithm (TLB): The calculation of TLB is appropriated VMs for controlling a specific occupation. The task manager has executed of each VM, utilizes the order and allocates the specified work given by the user. According to request of user, if task is fine for specific VM on the data center, then task is allotted to appropriate VM. If there no VM is accessible to send acknowledgement to the tasks, the task manager create a queue of requests [16].

## 7 Algorithms Implementation

Cloud is a novel technique in Information Technology (IT) environments with immense infrastructure and resources. The integral aspect of cloud is load balancing. Novel load balancing in cloud computing ensures effective resource utilization. Load balancer has two categories: the static load balancer and the dynamic load balancer.

### 7.1 DSBP Algorithm

To simulate the existing algorithms and the proposed DSBP algorithm for DDC, the algorithm is set as other algorithm in the configuration as shown in Fig. 2. All the five Data centers will be running using the DSBP algorithm. The simulation will provide the response time and the processing time cost at the DDC.



**Fig. 2** Performance of DSBP algorithm in DDC

**Table 1** Data center request servicing times in DSBP for DDC

DDC	Avg (ms)	Min (ms)	Max (ms)
Response time	46.48	0.00	53.40
Data center processing time	1.32	0.00	10.10

By running the simulation on Cloud Analyst tool, the tool generates the overall response time to process the request. The tool gives the minimum average and maximum time consumed by DDC to execute the requests and overall time consumed by system to handle the request. The maximum time, minimum time and average time are shown in Table 1.

Total virtual machine cost (\$): 0.30.

Total data transfer cost (\$): 0.07.

Grand total (\$): 0.40 (Table 2).

The implementation of all the algorithms; round robin, active monitoring, throttled load balancing and proposed DSBP algorithm. The DSBP takes advantages of all existing algorithms and achieve the efficient performance in response time and data transfer cost in centralized and distributed data center. So it is summarized that the results and findings which have been derived using the simulator i.e. Cloud Analyst for cloud based algorithms proved to be highly suitable and efficient for LIS in DDC. According to the all the implemented result, DSBP is far better for CDC and DDC to receive minimum response time and data transfer cost (Table 3).

**Table 2** Data transfer cost in DSBP for DDC

DDC	VM cost \$	Data transfer cost \$	Total \$
DC5	0.40	0.00	0.50
DC4	0.40	0.00	0.50
DC3	0.10	0.40	0.40
DC2	0.50	0.00	0.50
DC1	0.40	0.00	0.40

**Table 3** Simulation result for DDC

		Centralized data center			
DSBP	0.79	RT	151.40	0.12	160.10
		DCPT	0.30	0.12	0.80
Distributed data center					
DSBP	0.42	RT	46.48	0.00	53.40
		DCPT	1.32	0.00	10.10

## 8 Conclusion

Cloud Computing is an advances field and provides a systematic way to business or operations can operate in logistics management. Load balancing is one of key challenge to manage user requests on different data centers between the logistics users and partners in cloud computing. DSBP gives better results in both data centers environment over the cloud. As a matter of fact DSBP basically minimizes network congestion, higher latency and cost so that the various problems of multiple demand of Service in LIS can be efficiently solved. So, the DSBP algorithm has been implemented by using Cloud Analyst simulator kit which in turn finally analyze the result in comparison basis.

## References

1. K.-C. Shang, P.B. Marlow, Logistics capability and performance in Taiwan's major manufacturing firms. *Transp. Res. Part E: Logist. Transp. Rev.* **41**, 217–234 (2005)
2. M.S. Pishvaee, F. Jolai, J. Razmi, A stochastic optimization model for integrated forward/reverse logistics network design. *J. Manuf. Syst.* **28**, 107–114 (2009)
3. H.I. Hasiao, R.G.M. Kemp, J.G.A.J. Van Der Vorst, S.w.F. Omta, A classification of logistics outsourcing levels and their impact on service performance: evidence from the food processing industry. *Int. J. Prod. Econ.* **124**, 75–86 (2010)
4. Y. Kim, T.Y. Choi, T. Yan, K. Dooley, Structural investigation of supply network: a social network analysis approach. *J. Oper. Manag.* **29**, 194–211 (2011)
5. G. Buyukozkan, O. Feyzioglu, L. Bingol, Analytical Network Process for logistics management: a case study in small electronic appliance manufacturer. *Comput. Ind. Eng.* **58**, 432–441 (2008)
6. P.A. Horvath, C.W. Autry, W.E. Wilcox, Liquidity implications of reverse logistics for etailers: a Markov chain approach. *J. Prod. Retail.* **81**, 191–203 (2005)
7. Celebi et al., Analytical network process for logistics management: a case study in a small electronic appliances manufacturer. *Comput. Ind. Eng.* **58**, 432–441 (2010)
8. E.J. Bardi, T.S. Raghunathan, P.K. Bagchi, Logistics information system: the strength role of top management. *J. Bus. Logist.* **15**(1), 71–85 (1994)
9. S. Dubey, S. Jain, Logistics information system and cloud computing. *Res. J. J. Comput. Sci.* **1**(1) (2014)
10. V. Ilin, D. Simic, V. Svircevic, N. Saulic, Cloud computing applications in global logistics information system infrastructure. *J. Traffic Transp. Eng.* **1**(1, Serial No. 1), 30–38 (2016). ISSN 2328-2142
11. V. Tilokavichai, P. Sophatsathit, A. Chandrachai, Analysis of linkages between logistics information systems and logistics performance management under uncertainty. *Eur. J. Bus. Manag.* **4**(9) (2012). ISSN 2222-1905 (Paper)/ISSN 2222-2839
12. D.C. Marinescu, *Cloud Computing: Theory and Practice* (Morgan Kaufmann Elsevier, Waltham USA, 2013)
13. A.M. Salem, Facebook Distributed System Case Study for Distributed inside Facebook data center. *Int. J. Technol. Enhanc. Eng. Res.* **2**(7) (2017). ISSN: 2347-4289
14. S. Surpreeth, S. Birada, Scheduling virtual machines for load balancing in cloud computing platform. *Int. J. Sci. Res. (IJSR)* **2**(6) (2013). ISSN: 2319-7064
15. R. Nema, S.T. Edwin, A new efficient VM load balancing algorithm for cloud computing environment. *Int. J. Latest Res. Eng. Technol. (IJLRET)* **2**(12), 69–75 (2013). ISSN: 2454-5031

16. V. Asha, B. Kumar, V. Girish, Load balancing in cloud computing. *Int. J. Recent Trends Eng. Res.* **4**(3) (2018). ISSN: 2455-1457

# Mining Plants Features for Disease Detection Tensor Flow: A Boon to Agriculture



Saksham Goyal, Madhulika Bhatia, Urvashi, and Praveen Kumar

**Abstract** The abstract should summarize the contents of the paper and should contain at least 70 and at most 150 words. It's more difficult, complicated, time consuming and expensive to restore the health of a plant than to prevent illness. Tensorflow and Keras are considered to be the most important and useful libraries for a number of machine learning tasks. In this research paper we have used these two-machine learning techniques for the image mining of diseased plants. These are user friendly and efficient in nature. Lately, Handling of diseases using the traditional methods are laborious, time consuming as well as expensive. Since Agriculture is the primary occupation of 42.38% of Indians, managing of different types of plants and crops becomes very difficult due to the different diseases which arise in them. An estimate of 15–20% of the crops is destroyed in India every year due to pests, diseases and weeds. It is very important to develop an easy and convenient way of identifying these diseases so that it can prevent the loss of crops. This problem can be solved with machine learning technologies. This research paper introduces a novel technique to identify a plant's health status with the help of image classification and deep learning algorithms. The dataset is extracted as New Plant Disease data set containing healthy and unhealthy data of plant images. This study evaluates the images of plants taken as input and segregate different species by classifying them into healthy and unhealthy plants. The study uses the mathematical functions to identify dimensions and performs quantitative analysis of the images. The results are observed by evaluating on different metrics with training and validation accuracy as well as loss through visualizations to produce the final outcome as diseased or healthy plants. Train loss: “**–0.204**”, Train accuracy: “**0.03**”, Validation loss: “**–0.112**”, Validation accuracy: “**0.0312**”.

---

S. Goyal (✉) · Urvashi

Department of Computer Science and Engineering, Amity University, Noida, U.P. 201301, India  
e-mail: [goyal.saksham2000@gmail.com](mailto:goyal.saksham2000@gmail.com)

M. Bhatia

Department of Information Technology, Amity University, Noida, U.P. 201301, India

P. Kumar

Amity University Tashkent, Tashkent, Uzbekistan

**Keywords** Mining · Plants · Disease · Features · Tensorflow

## 1 Introduction

India is an agricultural country, wherein the entire population is dependent on this particular sector for its survival. There are a variety of factors which affect the crop yield that are soil fertility, availability of water, climate, and diseases or pests. Most of these factors are uncertain and completely depend upon the environment and area where the crops are getting cultivated. But, one of the major factors which affect the health of a crop are diseases [1]. As agriculture struggles to support the growing population, plant disease reduces the quality and quantity of the food, fiber and bio fuel crops.

This research paper will be discussing about a machine learning technique which will help this study with the problem. This study involves image classification of plants via Tensorflow and Keras [2]. With this process, one can classify an image on the basis of its visual content. Independent of the approach, identifying the health status of a plant correctly, in the starting stages is a crucial step for efficient disease detection.

In order to develop accurate image classifiers for the purpose of identifying the health status, there's a need of large, verified dataset image of unhealthy as well as healthy plants. The dataset is extracted, processed, analyzed before classification can be done using deep learning algorithms. The detailed framework is discussed in the paper and a complete model summary is explained with the feature extraction and detection process. The graphs showing the training and validation accuracy with its validation loss is determined using visualization tools.

## 2 Background

A Number of methods of image processing technique are used for the process of segmentation of plant leaf disease diagnosis. The challenging aspect is to categorize numerous types of diseases on various crops. Normally, red, green, and blue are the three colors for colored images. A lot of previous researches come up with errors when the application is implemented through RGB model. The reason is that, at that time the range varies from 0 to 255. For this purpose, RGB images are converted into grey images and in order to get the result, result histogram is estimated.

The most common type of diseases is Leaf spot, Brown spot, Mosaic, Grey spot and Rust. These diseases affect the apple field very badly.

## 2.1 Motivation and Contribution

A disease in plants causes a major loss to both the quantity and quality of the crops. These diseases have a great impact on the economic condition of a country and are a major threat to the food security. Farmers often experience great difficulty while dealing with diseases because it becomes very hard to find out if the plant is healthy or not with naked eyes. Farmers in India have hard times while dealing with plant diseases, one of the major reasons is lack of information and lack of understanding of the poor farmers and therefore, their disease management technique is very ineffective. Thus, we have to pay huge losses for the same every year. Prior information on the health status of a plant will help to take more effective measures to prevent diseases. Inappropriate preventive measures applied consistently in the production of crops will affect the food needs of the growing population that in turn would have serious economic effects, resulting in production losses and market declines.

## 2.2 Related Work

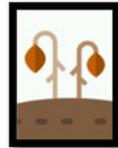
Jhuria et al. [3] used image processing for the detection of disease as well as fruit grading. Detection of diseases was done with the help of Artificial Neural Network. Initially, they created two autonomous databases, one for training images which already has stored images and other for the implementation of query images. In this, back propagation adjusts the weight of training databases. When the final results got calculated, they found out that the morphological feature gave us a better result as compared to the other two.

Sollapure [4] has done is that they have “Plant Leaf Disease Detection Based on Image Processing Using MATLAB “The plant leaf diseases are affecting the productivity quality and efficiency of the plants, thus effecting the growth of the country so that to overcome this plant leaf diseases must be recognized during initial stages. The proposed scheme is based on the image processing technique using MATLAB. Earlier linear support vector machine was used to identify single diseases. The machine learning methods bring this aspect to reality, by observing the database and helping the botanists in diagnosis of diseases where lot of precision is required. And one of the machine learning technique, SVM is used in this project for classification of leaf diseases.

## 3 Proposed Methodology

It can take a single image of the plant and it will identify whether the plant is healthy or unhealthy. If the accuracy is between (-1 and 0), then it is a Dead Plant, if the

**Table 1** Health status of plants

<b>-1 -0</b>	
<b>0-0.33</b>	
<b>0.33-0.66</b>	
<b>0.66-1</b>	

accuracy ranges from (0–0.33), then the plant is Unhealthy Plant, if it ranges from (0.33–0.66), it is Moderately Healthy Plant and if it is from (0.661), it is Very Healthy Plant as shown in Table 1. After this, with the help of accuracy graph, this study will further be working to diagnose an unhealthy plant and get necessary information from it. If the disease diagnosis based on automated image recognition, if is technically feasible, can be made available at lower levels. Table 1 shows Health status of plants with accuracy and loss values.

The dataset consists of about 87 K rgb images of healthy and diseased crop leaves which are divided into 38 different classes. The total dataset is divided into 80/20 ratio of training and validation set preserving the directory structure [5]. A new directory containing 33 test images is created later for prediction purpose.

### Algorithm used

An algorithm is being used by the Plant Disease Identification in order to execute all process for scanning the plants for the genuine prediction of diseases along with medical expertise supporting it.

**Step 1:** Setting up environment: The First important step is to set up and create an environment along with the user providing the dataset and access to the images.

**Step 2:** learning model: In Tensor Flow, existing algorithms learning model were used and is created on plants dataset.

**Step 3:** Training and Testing: Clustering algorithm is applied over data set and model is trained on plants dataset.

**Step 4:** Testing Data: Disease identification is a research objective of proposed work which allows the system to forecast plants disease on the sample of the new plant, on the basis of the various parameters provided by the farmers related to the plants if required.

**Step 5:** Optimization: Model optimization is applied in Tensor flow.

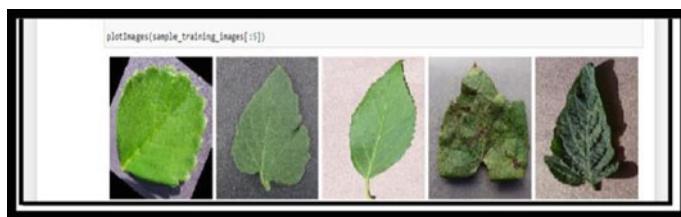
**Step 6:** Forecasting: This step is basically the prediction step where the user gets to know about the current situation of the event. This step also displays the history and the present advancement in the active diseases in order to enable the user tracking of the health process.

Data Extraction uses the ImageDataGenerator from the imported libraries from python, dense and other layers help in the finding out the health status of a plant in further research [6]. Next is Data Preparation.

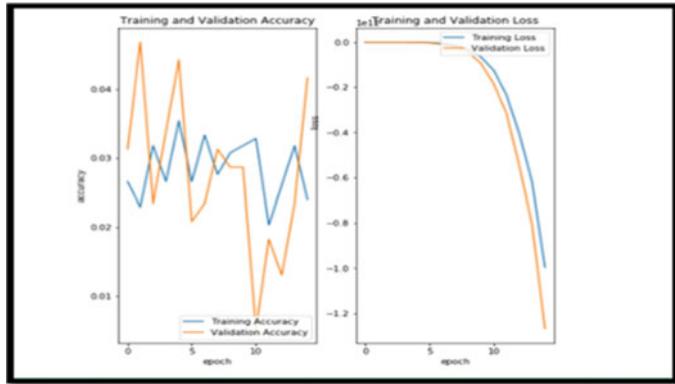
Data preparation needs the configuration of the images into suitable pre-processed floating-point tensors before delivering it to the network [7]. The task of transforming the images into pre-processed floating-point tensors requires analyzing the images from the disk and comprehends contents of these images into appropriate grid pattern as per their RGB content. Transforming them into floating point tensors and rescaling the tensors from values 0 and 255 to 0 and 1, the reason being that neural networks prefer to deal with small numbers like 0 and 1. Fortunately, all of these functions can be performed by the Image data generator class provided by keras.tf. Once the preparation of data is done, the URL for dataset which is stored inside the D disk needs to be entered on Tensorflow. The next step requires extracting of zip file and then extracting the dataset location [8].

The next step is to conceptualize the training images by withdrawing a batch of images from the training generator. Further, the requirement is to write a function that will plot images in the form of a pattern, the dimensions of the grid depend completely on the user. Here, the dimensions taken are 1 row and 5 columns. Printing of 5 images as a sample is done as displayed in Fig. 1.

Now, moving forward, create the model which consists of 3 convolution blocks with maximum pool layer in each of them. There is a fully connected layer with 512 units on top of it that is activated by a relu activation function. After this, compiling the model, view all the layers of the network using the model's summary method



**Fig. 1** Printing images from dataset



**Fig. 2** Accuracy and loss

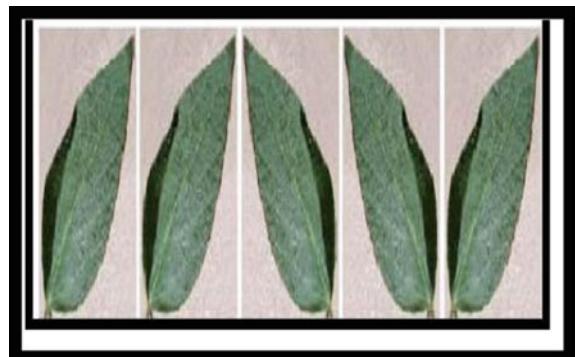
sing the `fit_generator` method of the `ImageDataGenerator` class to train the network. Parameters and optimization are done for achieving high performance training accuracy and validation accuracy [9]. Two parameters are taken into consideration for the above give task as shown in Fig. 2.

### 3.1 Data Augmentation

Over fitting in general occurs when there are a small number of training examples. One way to fix this problem is to augment the dataset so that it has a sufficient number of training examples [10]. Data augmentation is basically a simple process wherein more training data is generated from existing training samples by augmenting the samples used. Augmentation is done by working on random transformation that yields believable-looking images [11]. The goal is the model will never see the exact same picture twice during training. This helps expose the model to more aspects of the data and generalize better.”

Now, augment and visualize the data. Begin by applying random horizontal flip augmentation to the samples and analyze how individual images look after the transformation [12]. Begin by applying random horizontal flip augmentation to the dataset and see how individual images look like after the transformation as displayed in Fig. 3. Let’s take a look at a different augmentation called rotation and apply  $45^\circ$  of rotation randomly as displayed in Fig. 4.

Now, apply a zoom augmentation to the dataset to zoom the images by 50%. The zoom range is from 0 to 1 where 1 = 100%. Plotting the zoomed images randomly as displayed in Fig. 5. After that, applying all the previous augmentations [10, 13], here the following have been applied that are **rescale**,  **$45^\circ$  rotation**, **width shift**, **height shift**, **horizontal shift** and **zoom augmentation** to the training images. Visualizing

**Fig. 3** Horizontal flip**Fig. 4** 45° randomly rotation**Fig. 5** Zoom augmentation

how a single image would look after passing different augmentations to these images in Fig. 6.

**Fig. 6** Final augmentation

### 3.2 Dropout

Another technique to reduce over fitting is to introduce dropout to the network. It is a form of regularization that forces the weights in the network to take only small values, which makes the distribution of weight values more regular and the network can reduce over fitting on small training examples [14].

When applying 0.1 dropouts to a certain layer, it randomly kills 10% of the output units in each training epoch.

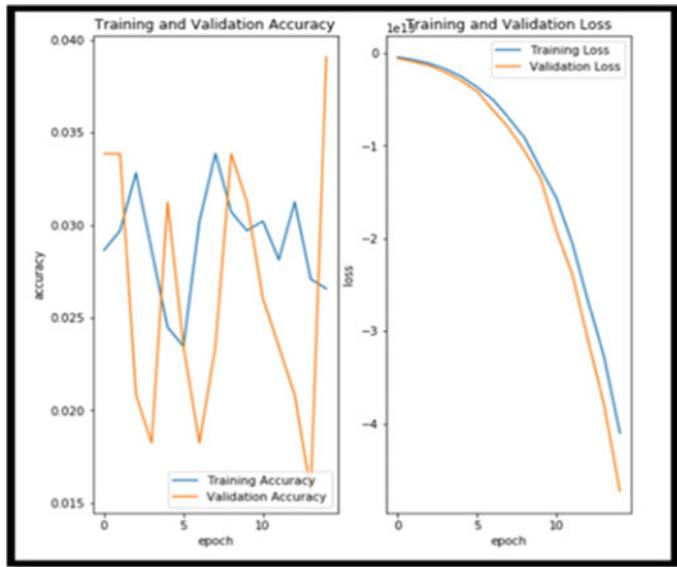
Creating a new network with dropouts, this helps in avoiding over fitting in the training dataset. Now, compiling the model. After successfully introducing data augmentation to the training examples and adding dropouts to the network, training this new network [15].

Now visualizing the model for better accuracy as displayed in Fig. 7.

## 4 Results and Discussions

### 4.1 Initial Observation

The first Training and Validation graph of accuracy shows a huge variation. The training accuracy was corrected (linear) since it was increasing linearly over a period of time whereas in the case of validation accuracy, it was not linear. This difference between the Training and Validation Accuracy means that the image was not analyzed properly because the picture uploaded was only from one angle (or view). So, to make the entire graph linear, dataset has to be augmented. Data augmentation refers to the creating of more training samples from the existing training samples by randomly transforming the image into different forms. All of these augmentations will help to have a better view of the dataset taken as shown in Fig. 2.



**Fig. 7** Accuracy and loss

After applying all augmentations of dataset, this is the final graph that is obtained and from the graph it is concluded that DATA AUGMENTATION helps in better accuracy as displayed in Fig. 7.

#### 4.2 Final Observation

Since the network was acquainted with only one single view, the training and validation accuracy were having a huge difference between them, whereas after augmenting the data, the network got acquainted with better view than before which assisted it give more accurate results. The graph of validation accuracy is now linear, which brings it to a conclusion that the accuracy graph is precise now. Therefore, the dataset of plant used is unhealthy in nature.

The overall observation from both the graphs is that whenever the graph is not precise and is giving vague results, one should opt for Data Augmentation because it will help to generate more data from the existing data.

Once the network is acquainted with data augmentation, it will become easier for the network to adapt to different architectures and amount of training data.

## 5 Conclusion

The main goal of this research paper is to help farmers focus on the initial stage of a crop while taking assistance of crop dataset and identifying the health status of a plant. Diagnosing the health status at the initial stage gives farmers a clear picture as to what possible precautions and measures have to be taken in order to protect the crop from further diseases. Our major aim was to reducing persistent monitoring of crops and produce high-yield results. The loss and Accuracy values of the dataset provide us with the correct status of a plant. The values of the dataset used in the research paper are

Train loss: “**-0.204**”, Train accuracy: “**0.03**”, Validation loss: “**-0.112**”, Validation accuracy: “**0.0312**”. With the help of Table 1, the result is that the dataset used is of unhealthy plants. Since, technology plays a crucial role in agriculture industry and diseases in crops have been a major issue for the farmers, this technique of image mining has a great future scope.

## 6 Future Scope

Advances in Image mining will benefit the farmers and in turn the government. Image mining in general is anticipated to help the farmers identify the disease and the treatment for it, Acknowledging and tracking people and things, diagnosing medical conditions, performing and accomplishing of surgery, reprogramming defects in human DNA. The future of image processing applications of satellite-based imaging ranges from planetary exploration to surveillance applications.

## 7 Limitations

This technology does come with some limitations:

1. It is expensive to some extent, based on system used and the number of detectors used.
2. One of the major limitations is that if the images are not clear, and they are smaller in size then even after data augmentation, one may not get efficient and precise results.
3. For uploading the databases, mining them, using the appropriate codes and then wait for the desired results is somehow time consumption, so for some extent it is also time consumer.

## References

1. G.K.P. Vyshnavi, M.R. Sirpa, M. Chandramoorthy, B. Padmapriya, Healthy and unhealthy plant leaf identification and classification using hierarchical clustering. *Int. Res. J. Eng. Technol. (IRJET)* **3**, 448–453 (2016)
2. V. Pallas, J.A. García, How do plant viruses induce disease? Interactions and interference with host components. *J. Gen. Virol.* **92**(12), 2691–2705 (2011)
3. M. Jhuria, A. Kumar, R. Borse, Image processing and fruit grading, in *Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*
4. R. Sollapure, Plant Leaf disease detection based on image processing using MATLAB. *Int. J. Innov. Res. Comput. Commun. Eng.* **6**(5) (2018)
5. A.H. Kulkarni, R.K.A. Patil, Applying image processing technique to detect plant diseases. *Int. J. Mod. Eng. Res.* **2**(5), 3661–3664 (2012)
6. S.P. Mohanty, D.P. Hughes, M. Salathé, Using deep learning for imagebased plant disease detection. *Front. Plant Sci.* **7** (2016)
7. S.A. Miller, F.D. Beed, C.L. Harmon, Plant disease diagnostic capabilities and networks. *Annu. Rev. Psychopathol.* **47**, 15–38 (2009)
8. A.N. Rathod, B. Tanawal, V. Shah, Image processing techniques for detection of leaf disease. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **3**(11) (2013)
9. N.R. Savaji, R.K. Rajbure, Plant Disease Detection Using Image Processing (2018)
10. A.M. Makowski, Feature extraction of diseased leaf images. *IEEE Trans. Inf. Theory* **59**(3) (2013)
11. P. Pawara et al., Data augmentation for plant classification, in *International Conference on Advanced Concepts for Intelligent Vision Systems* (Springer, Cham, 2017)
12. H. Beyene, Plant diseases prediction using image processing and machine learning techniques: survey. *Int. J. Comput. Appl.* (2250-1797) **1**(8) (2018)
13. H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik, Z. Al Rahamneh, Fast and accurate detection and classification of plant diseases. *Int. J. Comput. Appl.* (0975-8887) **17**(1) (2011)
14. S. Aggarwal et al., IOP Conf. Ser.: Mater. Sci. Eng. **1022**, 012118 (2021)
15. S. Aggarwal, M. Bhatia, R. Madaan, H.M. Pandey, SVM prediction model interface for plant contaminates. *Traitemet du Signal* **38**(4) (2021)
16. Z.B. Husin, A.H.B.A. Aziz, A.Y.B.M. Shakaff, R.B.S.M. Farook, Feasibility study on plants chili diseases using image processing Techniques, in *2012 Third International Conference on Intelligent Systems Modelling and Simulation*
17. R.A. Yeh et al., Semantic image inpainting with deep generative models, in *CVPR* (2017), pp. 5485–5493

# Data Visualisation Using Self-organising Maps



Madhulika Bhatia, Saru, Preeti Manani, Pratham Vats, and Praveen Kumar

**Abstract** The utilization of Self-Organizing Maps (SOM) to reprieve various knowledge data and complex datasets. Self-Organizing Maps helps in data training, getting error metrics, and convergence properties exist in SOM. Visualizing a Self-Organizing Map on data set create a visual difference in retrieving information from data. Self-Organizing Maps are capable to deliver the information regarding locations of the high dimensional groups, it also can support to catch irregular forms, and provide the understanding about the outline of the dataset applied on.

**Keywords** Self-organizing map (SOM) · Data · Information · Visuals

## 1 Introduction

In a multi dimension data there is a challenge to discover some or the other difficulties in examining data. There is a difference in training machine learning models on huge multidimensional datasets which inherents knowledge of the relationships and design to direct the training procedure and to help evaluate results. The Self-Organizing Map (SOM) is a dimensionality reduction method which gives perceptions about multi-dimensional data with nominal computing required. Self-Organizing Maps are used for empirical data analysis, complex clustering issues, and visualizing multi dimensional dataset. SOMs were established by Teuvo Kohonen in 1980 and have

---

M. Bhatia (✉) · P. Vats

Department of Computer Science and Engineering, Amity University, Noida, India  
e-mail: [mbhadauria@amity.edu](mailto:mbhadauria@amity.edu)

Saru

Department of Information Technology, Amity University, Noida, India

P. Manani

Dayalbagh Educational Institute, Agra, India

P. Kumar

Amity University Tashkent, Tashkent, Uzbekistan  
e-mail: [pkumar3@amity.edu](mailto:pkumar3@amity.edu)

contributed in many research papers and Journals. SOMs are created by mean of 2-D lattice of node.

A well trained SOM could be used to recognize how distinct data clusters allocated between a data region or how a reaction variable varies throughout the data set [1, 2]. For example if SOM and for every lattice node present in it, we map the mean output value of every data point closest to that particular node in the input region.

Research objective is to reinforce present knowledge about property relations using a data science method and an existing experimental data set of materials. This paper presents a Self-Organized Map (SOM) technique to create clustering and heat maps which will permit visual qualification of relations among material characteristics. SOM makes use of exploratory data to originate a strong visualizing tool which makes use of unproven machine learning to learn the knowledge of property relations of materials informatics [3]. On account of the visual output, SOM is an productive educational mechanism for teaching foundations of materials characteristics relationship to a larger audience. A complete workflow describing that enclosing data management, machine learning, and eventually, visualization to directly deduce the relationship linking materials [4]. A data science mechanism, to find a structured and practical method to acknowledge the fundamental inter-relations amongst the data which can potentially propose modern research directions [6–8]. SOM also constructs cluster maps which expose various unique materials present in the training dataset as a outcome of various processing techniques and complex species. Lastly, SOM is used to estimate new quantities, such as Grüneisen parameter, that were absent in our training data.

## 2 Methods for Visualizing High-Dimensional Data

### 2.1 Clustering Methods

Clustering is a form of unproven ML method. The inferences are drawn from a data collection which consist of unlabelled output variable in the unsupervised learning approach. It is an analytical data analysis methodology that helps to investigate the multivariable data sets [5].

Cluster is nothing besides the data point's company in a way so the separation linking the data points inside a cluster is minimum.

### 2.1.1 Types of Clustering Methods

#### Density-Based Clustering

In this procedure, the clusters are produced on the basis the compactness may be specified throughout the space. The sections that develop condensed because of vast amount of data points exist in that space are measured as clusters.

These area (the region where data points are existing in very less quantity) are measured as outliers. The clusters generated in these approaches can be of random profile. These are the instances of clustering algorithms recognized on the base of Density:

- DBSCAN (Density-Based Spatial Clustering of Applications with Noise) [9]
- OPTICS (Ordering Points to Identify Clustering Structure).

#### Hierarchical Clustering

Hierarchical Clustering sets also called Agglomerative groups because of distance metrics. In Agglomerative clustering, every data point behaves like a cluster in the beginning, and later groups the clusters one at a time.

These algorithms builds a distance matrix of which consists the existing clusters and make the linkage connecting the clusters relying because of link. The grouping of data points is characterized with help of dendrogram. Following are various types of linkages—

- Single Linkage;
- Complete Linkage.

#### Partitioning Clustering

This technique one of the widespread picks for analysts to build clusters. In partitioning clustering, the clusters are partitioned on the basis of properties of the data points. We need to describe the unit of clusters to be formed for the clustering technique. These clustering algos observe a repetitive procedure to re-assign the data points highly dependent on the distance. The algorithms which belongs to this class are:

**K-Means Clustering:** K-Means clustering is one of the extensively used algos. It separates the data points in k clusters on the basis of distance gathered for the clustering. The unit of ‘k’ has to be defined by user.

It is computationally costly algo as it calculates the distance of each data point from point of every cluster in every iteration [9].

## Grid-Based Clustering

The general technique in the algos of this procedure vary different algos. These algorithms are bothered with the value region adjoining the data points instead of the data points. One of the nest superiority of grid based algos is its depleting in computational complexity. This makes it perfect for handling huge data sets.

- STING (Statistical Information Grid Approach)
- WaveCluster

## 2.2 *Self-organizing Maps*

The SOMs were designed by Teuvo Kohonen as a exclusive form of neural networks. SOM is a model of particular feature of biological neural nets. SOM is a unsupervised machine learning model and adaptive recognition and representation strategy. SOM can define various features of a phenomenon in most of the domain, which provides the data in the domain to present with help of vectors of numerical features. The map trains by a self-organization procedure. The map comprises of a regular multi-dimensional (rectangular) grid of processing units—the neurons. Individual unit has an related model of some multidimensional observation, characterized as a vector of feature domain's value.

### 2.2.1 Properties Used in Analysing Data

By integrity of Self-Organizing Map's learning algorithms set up a nonlinear regression of reference vectors inside the input space. The reference vectors forms a multi-dimensional flexible network which follow the circulation of the figures.

**Ordered display:** The ordered character of the regression justifies using of the map as a presentation for data sets. Data is mapped to the scales on the map which has the nearest reference vectors, nearby values will be having same data items mapped onto them. One rich  $\beta$  benefit of every time utilizing the similar display is that as the predictors grow more confidential with respect to the map, they can infer new information demonstrated on it quicker and more simply.

**Missing data:** Most commonly taking place problem in applying procedures of statistics is that of missing data. Some of the constituent of the data vectors can't be accessible for most of data items, or may not even be relevant or well-defined. In the instance if Self-Organizing Maps the issue with absence of data can be handeled as follows: when selecting the winning unit by I vector  $x$  can be equated with the reference vectors  $m_i$  by means of only those modules that are present in  $x$ . Note that not a single reference vector components is missing. If only a minor share of

the components of the data vector is absent, the result of the evaluation will be statistically fairly precise.

### 3 Steps in Data Analysis Using SOM

#### 3.1 *Preprocessing*

Most of the unsupervised procedures simply demonstrate some structures present in data set, and the structures are eventually regulated by the features selected to characterize the data items. The utility of various preprocessing techniques rests strongly on its application. Generally, the function of tailoring involves substantial expertise both in the submission data analysis procedure. Some experiments which is possibly helpful as an automatic feature extraction stage are described in many publications. A system which takes learning from the context of words to characterize the words, and appropriately processed word category histograms to represent documents is also discussed in many Publications [10].

#### 3.2 *Computation of Maps*

Detailed procedures for how to essentially compute the maps are invented by Kohonen. The reference vectors initialized to lie in an ordered layout level traversed by the two main data. The preliminary segment imposes a global collection of the map, while in the subsequent phase the final precise state of the map is modeled gradually.

#### 3.3 *Selection of Maps*

Function to calculate cost for SOM is precise as the size of map anatomy of the map, which we get to know about through neighborhood kernel. The cost function's value will gradually decreases when we see an increase in map size, and increase whenever increase in the neighborhood function's width. Thus, it is not advisable to use it to compare maps with difference in sizes or neighborhood kernels; some secondary criteria are therefore needed. Topology salvation has, hence, resulted to be very challenging for determining a sensibly discrete grid. We basically have majorly two different ways for calculating the topology preservation degree. Some more points to be careful about are, are given below: **Scarce data**. If the data set is scarce, which will consist only few example per maps, this results in more difficulty in the process of measuring the topology preservation. After that probably it will be considered to be

useful, the distance to the third-closest reference vector etc., weighted properly. **High dimensionality.** Reason for contemplating the index of constancy to plotting starting first with input space to the map grid as a measure of the efficiency of the mapping stems because of the point that the whenever a SOM grid having a lower-dimension will automatically fold when it will follow a higher-dimensional distribution. Spaces in between along the mapping might just give us a signal of the presence of similar folds. We might have very less examples for differentiating a non-linear curve and an example from a higher-dimensional variation, for illustration. **Computational complication.** The planned degree is properly computationally concentrated. This process involves the way path in between each pair of units map, according to what distance value is from the reference of space vectors. The measure can, hence, be calculated using dynamic programming, which gradually decreases the difficulty. A rough guess of the measure, apt for very large maps, might also be attained by calculating the space alongside position vector of units which are there in shortest path of map frame.

### **3.4 Practical Implementation of SOM**

- **Interpretation.** The interpretation of the results of investigative data study depends, of course, on the application. Various general methods are present that might help during the process of interpretation. Need of proper careful examination is required during the atmost starting of the interpretation though the map image may be beneficial for automatically understanding about what all type of platforms will be useful.

## **4 Case Study**

For practically implementing of SOM, Fraud Detection dataset is used. The process that we followed is as follows:

This dataset that holds information about customers of a bank applying for a credit card and the customers provide at time of applying for the card to the bank authority. Goal is to detect the potential fraud application with the help of Self-Organizing Maps. The list of customers that potentially cheated. A supervised model to detect potential cheater is not used but instead use of unsupervised deep leaning method is applied i.e., using unlabelled data and working on them. Some patterns in identifying high dimensional datasets having non-linear relationships. One of these patterns would be potential fraud. The essential libraries are imported and then credit card application datasets is loaded. The following dataset has been extracted from UCI Machine learning repository- archive.ics.uci.edu/ml/datasets. The file contains credit card applications. All attributes' names also the respective value has been modified to meaningless symbols as shown in Fig. 1.

dataset - DataFrame																	
Index	CustomerID	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	Class	
0	15776156	1	22.88	11.46	2	4	4	1.585	0	0	0	1	2	100	1213	0	
1	15739548	0	22.67	7	2	8	4	0.165	0	0	0	0	2	160	1	0	
2	15662854	0	29.58	1.75	1	4	4	1.25	0	0	0	1	2	280	1	0	
3	15687688	0	21.67	11.5	1	5	3	0	1	1	11	1	2	0	1	1	
4	15715750	1	20.17	8.17	2	6	4	1.96	1	1	14	0	2	60	159	1	
5	15571121	0	15.83	0.585	2	8	8	1.5	1	1	2	0	2	100	1	1	
6	15726466	1	17.42	6.5	2	3	4	0.125	0	0	0	0	2	60	101	0	
7	15660390	0	58.67	4.46	2	11	8	3.04	1	1	6	0	2	43	561	1	
8	15663942	1	27.83	1	1	2	8	3	0	0	0	0	2	176	538	0	
9	15638610	0	55.75	7.08	2	4	8	6.75	1	1	3	1	2	100	51	0	
10	15644446	1	33.5	1.75	2	14	8	4.5	1	1	4	1	2	253	858	1	
11	15585892	1	41.42	5	2	11	8	5	1	1	6	1	2	470	1	1	
12	15500000	1	20.67	1.75	1	0	0	1.77e-1	1	1	2	1	0	140	211	0	

**Fig. 1** Data set loaded

Here, lines correspond to customers and columns corresponds to attributes. Customer segmentation is performed and identify segment of customer that potentially cheated. These customers are the inputs. These inputs are mapped to a new output space and between input space. Each neuron being initialized as vector of weights that is of the same size as vector of customer i.e. vector of 15 elements (customer id + 14 observation). This neuron is winning neuron. For each customer the winning node is the most similar neuron to the customer.

Outliers are in the region It would show to detect outliers and hence detect frauds.

The inverse mapping functions is used to identify which customer originally in input space associated with winning node that. Two subsets are created. X contains all the variables except the last one as shown in Y contains the last variable that tells yes or no whether the application is approved or not in Figs. 2 and 3.

The data set split into x and y. It will predict zero or one at the last. Distinction are retrieved that the customers who are approved or not. Feature scaling is done because there are some high computations to be made as shown in Fig. 4.

Weight vectors are initialized to small numbers then self-organizing maps are trained on data set x. The method will be repeated. The results are outline with neurons inside the maps. Two-dimensional grid, that contains all the final winning nodes, and for each of the winning nodes we will get MID (Mean interneuron distance). MID of specific winning node is the mean of the length of all the neurons around the winning node. Higher the mid more the winning node will be far away from the neighbours. Higher is the mid the more the winning node is an outlier Mostly all the winning nodes represent the rules which are cherished. These outliers represent the frauds. It will be done by using colours. The winning nodes will be coloured by different

Y - NumPy object array	
0	0
1	0
2	0
3	1
4	1
5	1
6	0
7	1
8	0
9	0
10	1

**Fig. 2** Subsets X

X - NumPy object array															
0	1.57762e+07	1	22.08	11.46	2	4	4	1.585	0	0	0	1	2	100	1213
1	1.57395e+07	0	22.67	7	2	8	4	0.165	0	0	0	0	2	160	1
2	1.56629e+07	0	29.58	1.75	1	4	4	1.25	0	0	0	1	2	280	1
3	1.56877e+07	0	21.67	11.5	1	5	3	0	1	1	11	1	2	0	1
4	1.57158e+07	1	20.17	8.17	2	6	4	1.96	1	1	14	0	2	60	159
5	1.55711e+07	0	15.83	0.585	2	8	8	1.5	1	1	2	0	2	100	1
6	1.57265e+07	1	17.42	6.5	2	3	4	0.125	0	0	0	0	2	60	101
7	1.56604e+07	0	58.67	4.46	2	11	8	3.04	1	1	6	0	2	43	561
8	1.56639e+07	1	27.83	1	1	2	8	3	0	0	0	0	2	176	538
9	1.56386e+07	0	55.75	7.08	2	4	8	6.75	1	1	3	1	2	100	51
10	1.56444e+07	1	33.5	1.75	2	14	8	4.5	1	1	4	1	2	253	858

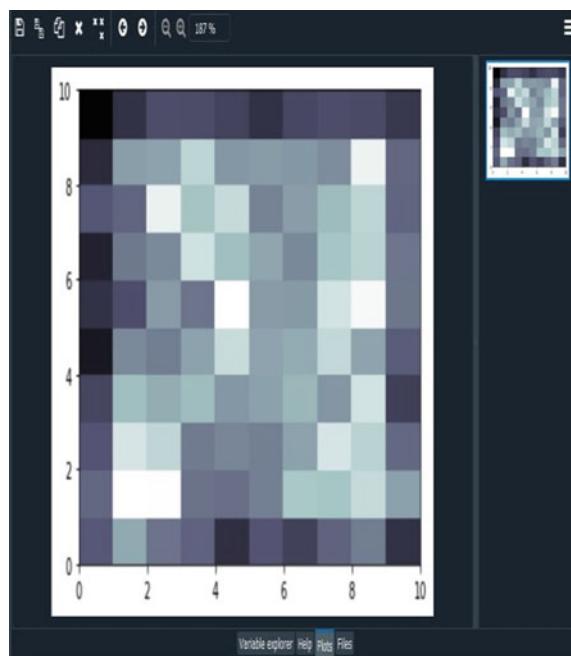
**Fig. 3** Subsets Y

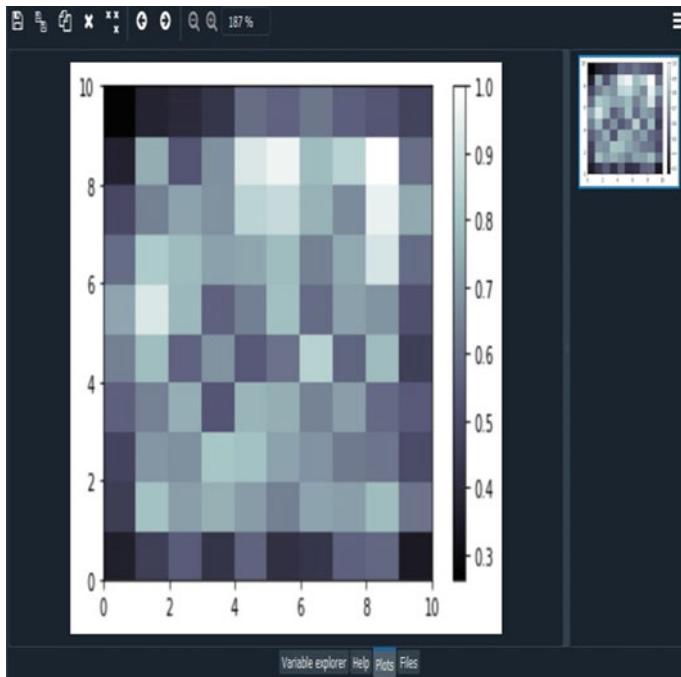
X - NumPy object array							
	0	1	2	3	4	5	6
0	0.842681	1	0.125263	0.409286	0.5	0.230769	0.375
1	0.696091	0	0.134135	0.25	0.5	0.538462	0.375
2	0.388982	0	0.238045	0.0625	0	0.230769	0.375
3	0.488425	0	0.119098	0.410714	0	0.307692	0.25
4	0.600795	1	0.0965414	0.291786	0.5	0.384615	0.375
5	0.0216515	0	0.0312782	0.0208929	0.5	0.538462	0.875
6	0.643706	1	0.055188	0.232143	0.5	0.153846	0.375
7	0.379115	0	0.675489	0.159286	0.5	0.769231	0.875
8	0.393338	1	0.211729	0.0357143	0	0.0769231	0.875
9	0.2919	0	0.631579	0.252857	0.5	0.230769	0.875

**Fig. 4** Feature scaling

colours with respect to the path which is larger is the MID, the closer to white the colour will be in form of in Figs. 5 and 6.

**Fig. 5** MID matrix



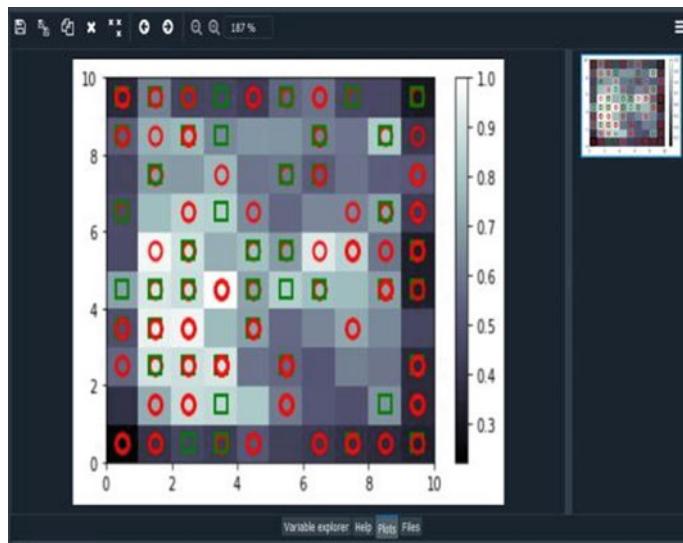


**Fig. 6** MID color

The white colour matches with a large MID and the dark colours corresponds to low mIDs. Legend on the right is the range of values of MID. These are normalised values that is the values are scaled from zero to 1. The high interneuron distances corresponds to the white colour. And on the prior side, the smallest interneuron distances correspond to the dark colour Frauds are identified from the outline winning node because this is different from the general rules. All the majority of points with dark colours are close to each other because their MID is less. Two markers showing red circles correspond to client who didn't get approval and green circles corresponds to the customers who got approval. A variable marker is there and create a vector of two elements corresponding to two elements. Red colour quoted by r and green colour quoted by g. Circle on the customers and for each customer, winning node is retrieved and depending whether the customer get approval. Iterator variables I and x. There is distinct value of all the indexes of record dataset x will be different vectors for records. Main node we have a method that returns the winning node for specific customer. (Winner). Each winning node is Coordinates of the winning node are ( $w, 0$ ) and ( $w, 1$ ). These are the winning node of lower left corner. Centre so will put + 0.5 to put it centre. The most important thing is to know whether we have to put red circle or green circle. Passing the markers vectors inside the plot functions and inside the brackets we will pass the y vector and the index I. Customer did not get approval  $y[i]$  will be zero and marker [0] will be circle. And if customer get approval  $y[i]$  will

be 1 and markers [1] which corresponds to square. Taking the colours vector and passing the y vector and index i. Colour the marker edge so pass the markered colour parameter. SOM which will depict customers associated with the winning nodes are customers who got approval or did not get approval as shown in Figs. 7 and 8.

Apart from MID we get information whether the customer gets approval or didn't get consent for each of the winning nodes. Customer related to the winning node



**Fig. 7** SOM

mappings - Dictionary (71 elements)		
Key =	Type	Size
(0, 0)	list	10
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(0, 2)	list	2
		[Numpy array, Numpy array]
(0, 3)	list	6
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(0, 4)	list	10
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(0, 6)	list	18
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(0, 8)	list	19
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(0, 9)	list	12
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(1, 0)	list	5
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array]
(1, 1)	list	4
		[Numpy array, Numpy array, Numpy array, Numpy array]
(1, 2)	list	48
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(1, 3)	list	8
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]
(1, 4)	list	6
		[Numpy array, Numpy array, Numpy array, Numpy array, Numpy array, Nump ...]

**Fig. 8** X dataset

Index	Type	Size	Value	...
7	Array of float64 (15,)	[0.97740351 1. 0.30571429 ... 0.5 0.2215 0. ...		
8	Array of float64 (15,)	[0.94645396 1. 0.19924812 ... 0.5 0.13 0.01065 ...		
9	Array of float64 (15,)	[0.90064029 1. 0.35714286 ... 0.5 0.082 0.004 ...		
10	Array of float64 (15,)	[0.73816417 1. 0.33954887 ... 0.5 0.1 0. ...		
11	Array of float64 (15,)	[0.45684322 1. 0.30451128 ... 0.5 0.12 0. ...		
12	Array of float64 (15,)	[0.73193342 1. 0.45984962 ... 0.5 0.16 0. ...		
13	Array of float64 (15,)	[0.66964189 1. 0.30827068 ... 0.5 0.0015 0. ...		
14	Array of float64 (15,)	[0.81905185 1. 0.2756391 ... 0.5 0.18 0. ...		
15	Array of float64 (15,)	[0.83320319 1. 0.35593985 ... 0.5 0.2 0.058 ...		
16	Array of float64 (15,)	[0.91761069 1. 0.31834586 ... 0.5 0.1195 0.002 ...		
17	Array of float64 (15,)	[0.55977448 1. 0.18676692 ... 0.5 0. 0. ...		
18	Array of float64 (15,)	[0.89896981 1. 0.26796092 ... 0. 0.464 0. ...		

**Fig. 9** Winning node

with red circle didn't get consent but the customer. Associated with winning node having green circle got the approval. Colour is around this range so it did not indicate high risk of fraud. If we look at the outlets (Winning node with white colour). Here the mean interneuron distance (MID) is almost equal to 1. Which clearly indicates there is high risk of fraud among the customers associated.

In this winning node we see that we have both cases, the customers who got approval. Map to catch these potential cheaters `Som.win_map(x)`. Here the key is the co-ordinates of the winning nodes. For this winning node we get the list of all records related to the winning node. It can be witnessed that there are 19 Customers in this list associated with this winning node as shown in Figs. 9 and 10.

Each line corresponds to one customer that is associated with these winning nodes of coordinates 0, 8. By clicking on the customer will get a window that displays the values of various attributes. This mapping to detect the fraud. Declaring a variable `fraud`. We get the coordinates of the winning nodes which had colour white. By passing the coordinates of this winning node to the mappings i.e., by doing the slicing with coordinates (3, 4). This will give us the list of customers associated with this outline winning node. These values are normalized. We will use inverse mapping function to get the original scales. Second outline winning node is with the coordinates (2, 3).

frauds - NumPy object array

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
11	1.57636e+07	0	16	0.165	2	6	4	1	0	1	2	1	2	320	2
12	1.56622e+07	0	28.58	3.75	2	8	4	0.25	0	1	1	1	2	40	155
13	1.56964e+07	0	31.57	1.5	2	1	1	0	0	1	2	1	2	200	106
14	1.55685e+07	0	15.83	7.625	2	11	4	0.125	0	1	1	1	2	0	161
15	1.57822e+07	0	22.67	0.75	2	8	4	2	0	1	2	1	2	200	395
16	1.55865e+07	0	27.25	0.29	2	7	8	0.125	0	1	1	1	2	272	109
17	1.57177e+07	0	23.58	0.83	2	11	4	0.415	0	1	1	1	2	200	12
18	1.57054e+07	0	30.58	10.665	2	11	8	0.085	0	1	12	1	2	129	4
19	1.57501e+07	0	25.25	13.5	1	1	1	2	0	1	1	1	2	200	2
20	1.56465e+07	0	17.92	0.54	2	8	4	1.75	0	1	1	1	2	80	6
21	1.575e+07	0	27.58	3	2	7	4	2.79	0	1	1	1	2	200	11

Format    Resize     Background color

Save and Close    Close

**Fig. 10** Potential frauds and original values

## 5 Experimental Results

Our model accomplished certain accuracy fraud detection method and the concerned population was established to be 22.60%. The consequences will vary if the same jupyter notebook run as initialization of the weights of the nodes of SOM grid is finished by casually choosing the records/patterns from the input sector i.e. arbitrarily shortlisting the records among the given dataset. Future work will be on various iterations like 100, 200 etc. to have more precise convergence and accumulate the weights of the SOM so that we can attain improved accuracy.

## 6 Conclusion

As results demonstrates that SOM provides improved outcomes in instance of detection of credit card fraud. As all limitations are confirmed and well presented in plots. The authenticity of our technique lies in making the use of normalization and clustering procedures of SOM of identifying credit card fraud. This will help in detection

of hidden pattern in the transactions that could not be identified with respect to some other traditional technique. With suitable number of weight neurons and by taking the help of thousands of iterations the network is trained and then outcome is tested to new transactions.

## References

1. J. Hollmen, *Self-Organizing Map (SOM)* (Aalto University, 1996)
2. H. Shah-Hosseini, R. Safabakhsh, TASOM: a new time adaptive self-organizing map. *IEEE Trans. Syst. Man Cybern. Part B: Cybern.* **33**(2), 271–282 (2003)
3. S. Jojo, S. Dhir, Mahurima, A machine learning APP for prediction of pandemic such as COVID-19, *Trends in Future Informatics and Emerging Technologies*, Advanced Computing Techniques: Implementation, Informatics and Emerging Technologies (Bentham Publishers, 2021), pp. 27–38. ISBN: 978-981-4998-45-1 (Online)
4. B. Everitt, *Cluster Analysis* (Wiley, Chichester, West Sussex, UK, 2011)
5. A. Shukla, S. Dhir, Tools for data visualization in business intelligence: case study using the Tool Qlikview, in *International Conference on Information System Design and Intelligent Application 2016, published proceeding in Advances in Intelligent Systems and Computing*, vol. 433, Springer Journal Volume 434, pp. 319–326
6. M. Vyomkesh, S. Dhir, A. Garg, Statistical analysis of Indian E-commerce market, in *International Conference on Soft Computing Techniques and Implementations, ICSCTA2015*, IEEE
7. A. Garg, M. Madhurima, S. Dhir, Hubbubs of research terminologies: let cat out of bag. *IJHEPS* **1**(9), 1–10 (2015). ISBN:978-81-909047-9-7
8. S. Gupta, S. Dhir, An enhanced approach to use SSL for end to end security. *Int. J. Comput. Sci. Inf. Technol.* **5**(2), 1053–1057 (2014). ISSN: 09759646
9. E. Bajal, V. Katara, M. Bhatia, M. Hooda, A review of clustering algorithms: comparison of DBSCAN and K-mean with oversampling and t-SNE. *Recent Patents Eng.* (2021)
10. M. Vashisht, M. Bhatia, Role of mathematics in image processing, in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, pp. 538–543. <https://doi.org/10.1109/COMITCon.2019.8862438>

# SQL Injection Attacks, Detection Techniques on Web Application Databases



Navdeep Singh and Preeti Tiwari

**Abstract** Most of web-application Databases are vulnerable to SQLi Query Injection Attacks, which lead to the entry of confidential data right through the client-side. They run by entering malicious SQL Injection Query codes through the web API on the client-side and return all the sensitive and private data from the database. SQL Injection is a technique where the web attackers post the malicious SQL injection Query with a specific end and goal there is to change the structure, behavior of the Query proposed by the computer programmer and occupying up the full admin login access of the web database, for malicious Input data modifications or deletion of the existing user User's Information. The SQLi attacking method is a critical web application attack; using this technique, attackers execute or run evil SQLi queries or an arbitrary code on the database servers using the web based application to gain fully take over the important and Private Data on the database bypassing authentication level. According to the owasp (Open Source Foundation for Application Security), the SQLi ("SQL Injection") attacks is mentioned as a significant Database Security Issue. The SQL injection query attack (SQLi) is the web-based vulnerability that allows the web attackers to spoof the identity, destroy the data presented on the database server, Insert new entries in the database and change the records submitted on the database. The significant issues of SQL Injection (SQLi) include loss of users' confidentiality and authentication problems, Authorization Issues for users, and loss of Data Integrity as any unauthorized use may change any sensitive information. In (SQLi) attack, there are three main types of SQL injection attack, i.e. In-band (classic (SQLi) Attack), inferential (blind(SQLi) attack), and Out-band SQLi Attack. In the Classic-SQLi Attack, the attackers use the same channel of communication that authorized user uses to launch attacks & extract the results. The Attack can generate errors on the database or use UNION operation to integrate multilevel Select statements to get only one single database as HTTP Response in web application. There are other procedures methods that researchers have proposed

---

N. Singh (✉) · P. Tiwari  
IIIM, Jaipur, India  
e-mail: [nsdhukia789@gmail.com](mailto:nsdhukia789@gmail.com)

P. Tiwari  
e-mail: [preeti.tiwari@icfia.org](mailto:preeti.tiwari@icfia.org)

to block or find SQL injection attack. That paper will include comparative study of advanced technique and tools to find SQL injection query attack (SQLi).

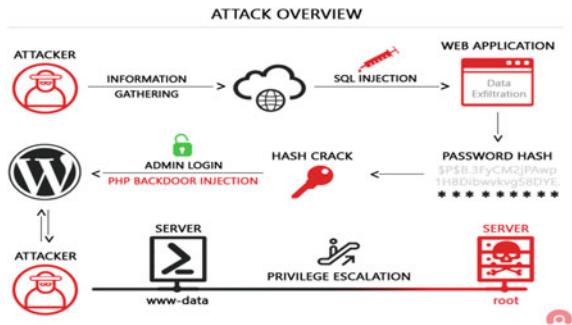
**Keywords** Web application database security · SQL injection · SQLi threat · Spoofing · SQL authentication bypass · Firewall bypass · SQLi attacks · SQLi poisoning

## 1 Introduction

Web Application Vulnerabilities in Web Database create major Data security and privacy problems like the disclosure of personal data, loss of User's faith, Sqli, and denial of service (DDoS). In line with OWASP [1], SQL query injection vulnerabilities are the foremost major vulnerabilities for Web application database systems. In the SQL query injection vulnerabilities: XPath query injection (XPathi), XML query injection (XMLi), and Postgre query injection vulnerabilities are mostly listed in Web applications and Web Database that use MySQL MsSQL, oracle, Postgre SQL databases. These SQLi vulnerabilities are usually prompted by the utilization of user inputs in security-sensitive program statements without proper authentication.

Nowadays, there are so many Web application databases that are unprotected from SQL injection attacks. For the reason that the character of the SQLi attack which is the unofficial entrance to the User's Private data and adding, deleting and, altering it, this type of SQLi attack is most well-liked among unauthorized Users, and again due to the indicated motive it is essential to assemble web-based applications unbreakable against SQL injection attacks. The SQLi attack Occurs when an unauthorized person attempts to access unauthorized entry to a web application database by entering evil SQLi query that modify the logic, data, syntax, and semantic of the authorized queries. And in return, there is a huge suggest technique to find the loopholes in the web-based applications and block the external and internal attacks. To be more particular, this observation pointing out some different-2 papers with various suggested techniques and automated tools to determine and stop the SQLi attacks from being able to get a huge scope of techniques in this way and differentiate them reverse each other. Here we explained an approach in this research to be able to get huge details from every one paper. So, after that, this is much uncomplicated to differentiate the technique and examine them. We concentrated on the important areas and stages of each technique suggested in each research paper, and then We found the limitations and delimitation of them at the end. And also, the Researcher discussed different types of SQL query injection attacks in that research paper.

**Fig. 1** Architecture of SQL injection attacks



## 2 Architecture and Framework for Security Vulnerabilities Proposed

According to a literature survey analysis of various vulnerabilities, SQL query injection attack is one of the important frequent security's flaws. Various academics have suggested a variety of technologies and mechanisms to counteract the SQL attack.

Our study discovered that SQL injection attacks are still feasible even after adopting the current market available prevention mechanisms; therefore, we suggested the architecture illustrated in Fig. 1 to offer a preventative method. We've suggested a technique in this diagram for attacking any website using SQL injection, regardless of whether any of the existing available safeguards block it.

All sessions are described in our architecture (Fig. 1).

**Users:** A web browser's user effectively makes web page requests from web servers all across the World Wide Web. The user logs into the system usually, make a requests to the server, and receives a responses in this article. This only occurs in a typical situation.

**Attackers:** An unauthorized user is referred to as an attacker (hacker). Essentially, this types of attackers would be a skilled programmers or engineers with adequate technical expertise to identify the unsecured spots and loopholes in a web application security systems. In that post, the attackers uses SQL query injection to assault the website. The uses of SQL injection attacks have already been discussed.

## 3 SQL Injection Attacks

SQL queries injection attack are often used to infect websites that enable user to put information into form field for databases queries. An SQLi is a malicious query command delivered to an interpreter that was not intended. The SQL query for user information may be changed, deleted, and locked by the web attackers. Data operations such as deletion, creation, update, and modification queries interact directly

**Table 1** SQL injection strings and its results

S. No.	Sql injection strings	Result obtained
1	1	First-name, sur-name
2	%'or'0' = '0	This cmd shows all records in db
3	%'or 0 = 0 union select null, version() #	This cmd shows the version of the db
4	%'or 0 = 0 union select null, user() #	This cmd shows the user of the db
5	%'or 0 = 0 union select null, database() #	This cmd shows the database name
6	%'or 1 = 0 union select null, table name from info_schema.tables #	This cmd shows the table name
7	%'or 1 = 0 union select null, table name from info_schema.tables where table name like 'user%'#	This cmd shows the user-name and passwords from tables

with databases. The queries establish a connection between the static component and the attack value (Table 1).

### Types of SQLi Query Attack: (1) Union-Based SQL Query Injection Attacks

By enlarge the results provided by the initial query, union-based SQLi enables an unauthorized user to collect data from the online database. Only put the Union operators if both the interrelated and fresh queries has original construction [1].

“This kind of attack may be carried out by targeting a UNION based query in the susceptible URL values, which reflecting a data-set that is the union of the enlarge first-one query’s result and the result of the SQLi injected query” [2]. “The SQL UNION operator joins the results of twice or grater searches to create a result set containing fetched rows data from the UNION’s participating queries” [3] (Table 2).

### (2) SQLi Error-Based Query Attacks

Error-based SQLi Query attacks, also known as in-band SQL injection, works with error messages produced by the database server and website in online applications to acquire information about the database’s structure and information. An unauthorized person may sometimes enumerate full database information only by using error-based SQL injection (Table 3).

**Table 2** Result set containing fetched rows data from the UNION’s participating queries

Description	Query
Union based SQLi	SELECT "mysql" UNION SELECT database()
Union subquery SQLi	SELECT "mysql" UNION (select database())
Union based null query	SELECT "mysql", "test" UNION SELECT @version,null
Stacked based SQL queries	SELECT "mysql"; INSERT INTO 'user' ('password') VALUES ((SELECT @version))

**Table 3** Results or to collect data on how to reconstruct the query for opposite injection

Description	Query
XML parse error with db version	SELECT extractvalue(rand(),group_concat(0 × 2a,(select @@version)))
Double query SQLi	SELECT1AND(SELECT1FROM(SELECT COUNT(*),group_concat(0 × 3a,(SELECT user_name FROM USER LIMIT 2,1),FLOOR(rand(0)*2))x FROM information_schema.TABLES GROUP BY x)a)

### (3) Blind-Based SQL Query Injection Attack

One of the important sophisticated SQL Query injection attacks is the blind SQL injection attack. Blind SQLi is divided into two categories. The first is a partial-blind method, while the second is a full-blind technique, both of which are described below. Use caution while running these queries, since they may overburden the server if they're run with a lot of automation.

#### Partial-Blind SQL Injection

SQL injections that return HTTP Status Code or another indicators in the HTML response that indicates true (1) or false (0) (1, 0) statement are known as partial-blind SQL injections. The queries in Table 4 will attempt to exploit SQL injection by asserting a true (1) or false (0) response statement based on assumed data. Returning True (1) or False (0) rows may also be used to identify true or false queries. An error message may also be utilized to determine if a value is 0 (False).

#### Full-Blind SQL Injection

Differences in HTTP status codes and contents in the HTTP response in the web application may be used to identify full-blind SQL injections.

Full-blind SQLi may also show whether a statement is true(1) or false(0). The table's queries will try to exploit the injection by declaring a true or false answer based on assumed data. Returning True(1) or False(0) rows may also be used to identify true or false queries. All database information is shown with an error response in full-blind SQLi.

**Table 4** SQL query and its results description

Description	Query
Version is 6.a.b.c	SELECT substring(@@version,1,2) = 5
If sub select is enabled	SELECT 1 AND (select 1) = 2
If Table data_table available	SELECT 1 AND (select 2 from data_table limit 3,2) = 2
First letter of first message is g	SELECT ascii(substring((SELECT message from data_table limit 3,2),1,1)) = 113

## 4 SQLi Attack Detection and Prevention Techniques

Many tools for detecting and preventing SQLi attacks have been developed for SQL injection. Ex: AMNESIA SQLi detection tools, created by Halfond and Orso, are SQLi attack detection and prevention tools. For this purpose, that tool is used for static analysis and runtime monitoring. Each hotspot's SQLi detection and prevention tool creates a model of valid queries. For detecting and preventing SQL injection, the SQLi tool CANDID has been suggested. That tools refer to the programmer's planned query structure dynamically and statically on any input data, data entry, and identify threats by comparing them to the intended query structure. To deceive the attacker, the SQLrand tool utilizes instruction set randomization to identify and abort SQLi queries containing injected code, and each SQL term is coupled with a random number. By integrating dynamic application code analysis with run-time validation, the proposed approach avoids SQLIA in stored procedures. Using a pattern-matching algorithm, an effective method for identifying and blocking SQLi attacks is described.

### (i) Sanitization

Validation and sanitization of data are critical preventive methods that must be carefully followed. Sanitization is the process of checking random data provided using the data entry field in the form of a functions is ensure that there are no dangerous or unnecessary characters in the SQLi query [4].

### (ii) Web-application firewalls

The web-application firewall is a piece of software or technology that many big businesses invest millions of money in to block or filter dangerous material. Web application firewall identifies and stops the attacker from using the database system to obtain access. A firewall is a software or hardware system that protects against Internet assaults [5].

### (iii) Using appropriate privileges

If an administrator-level dialog box prompts you to access the database, don't do so unless you have a compelling cause or it's required. As a result, attackers may simply get access to the user account. As a outcome, attempt to utilize an account with a locked set of capabilities that is always under the admins control [6].

### (iv) Stored-Procedure

Instead of manually entering data and commands, the user may call the stored-procedure using the web applications point of views, which is saved in to databases.

### (v) Produce Declaration

The entire SQLi code in which variable are given to it is defined by parameter SQL queries [7]. Related on this Sql query injection, the database-server is able to distinguish between sql-codes and information without having to worry about the entered information (Table 5).

**Table 5** Differentiation of SQL injection detection techniques

Tools	Tautology	Stored procedures	Blind attack	Illegal	Piggy backed	Union	Alternate encoding	Timing attack
Sqlrand	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Csse	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Candid	Yes	No	No	No	No	No	No	No
Fgl	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Proposed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

## 5 Proposed Method

The fundamental types of sqli attacks, as well as preventive and detection methods using encoding techniques such as the 2-way encoding method, are discussed in this article. This technique encodes the text before saving it in the database, for example, Base64 or hex code. This will assist to reduce the number of assaults. This makes the string unbreakable. Before they are executed, the SQL queries that include the data are stopped. Randomization of instruction sets aids in the encoding of SQL keywords. To decode them, a proxy server is utilized, and it filters requests that include clear-text phrases.

## 6 Conclusion

Web application programs are used by a large number of online companies and IT organizations nowadays to offer services to their customers. To provide accurate data, web applications rely on the back-end database connection. Data kept in databases, on the other hand, are often targeted by hackers or unauthorized users. SQLi is a standard method used by attackers to gain access to databases through SQL queries. The researcher examined several kinds of SQLi attacks, SQLi types, detection, and protection methods in the study. Basic SQL injection attacks and detection methods come in a variety of shapes and sizes, as the researcher said. The researcher was able to get a thorough understanding of the current SQL Injection Attack methods on web-based applications and database systems via this article. The study also looked at the many ways in which a SQL Injection attack may be implemented into a program. The researcher has also discovered and implemented methods for securing online applications and database systems against SQLi attacks.

## References

1. C. Huang, J. Liu, Y. Fang, Z. Zuo, A study on web security incidents in China by analysing vulnerability disclosure platforms. *Comput. Secur.* **58**, 47–62 (2016)
2. S.J. Sadana, N. Selam, Analysis of cross site scripting attack. *Proc. Int. J. Eng. Res. Appl. (IJERA)* **1**(4), 1764–1773 (2011)
3. S. Ali, S.K. Shahzad, H. Javed, SQLIPA. An authentication mechanism against SQL injection. *Eur. J. Sci. Res.* **38**(4), 604611 (2019)
4. L.S. Shar, H.B. K. Tan, L.C. Briand, Mining SQL injection and cross site scripting vulnerabilities using hybrid program analysis, in *Proceedings of the International Conference on Software Engineering (ICSE '13)* (IEEE Press, 2013), pp. 642–651
5. A. Avancini, M. Ceccato, Towards security testing with taint analysis and genetic algorithms, in *ICSE Workshop on Software Engineering for Secure Systems*, vol. 5 (2010), pp. 65–71
6. R. Johari, P. Sharma, A survey on web application vulnerabilities (SQLIA, XSS) exploitation and security engine for SQL injection, in *Communication Systems and Network Technologies (CSNT)* (IEEE, 2012), pp. 453–458
7. R. Kumar, Mitigating the authentication vulnerabilities in Web applications through security requirements. *Inf. Commun. Technol. (WICT)* **60**, 651–663 (2011)

# Deep Learning Based Brain Tumor Segmentation: Recent Updates



Rita B. Patil, Nirupama Ansingkar, and Prapti D. Deshmukh

**Abstract** This paper reviews various brain tumor segmentation techniques that used to recognize the detection of position and extension of tumor areas like lively tumorous tissue, sedative tissue, and effusion of serous fluid into the interstices of cells in tissue spaces or into body cavities. Dice score, Sensitivity and Specificity are three mostly used performance measures for the brain tumor segmentation. This paper focuses on existing literature in terms of the result of dice score performance measure for whole tumor, core tumor and active tumor segmentation of glioma brain tumor.

**Keywords** Magnetic resonance imaging (MRI) · Computed tomography (CT)

## 1 Introduction

The abandoned, unusual development and separation of the cells in the body known as cancer. The existence of this unnecessary cell development and separation in the brain tissue as a mass is said as brain tumor [1]. These tumors can give the impression at any place in the brain and have practically in any Size, contrast and shape according to their own nature. Thus it crates complexity in segmentation of brain tumor.

The meningioma type of tumors is able to effortlessly segmented whereas glioblastomas and gliomas are significantly complicated to segment and localize due to its surrounding edema which is often tentacle-like structure, diffused and poor contrasted. The kind of brain tumor which is mostly get in children is glioma, which can be low grade or high grade i.e. less aggressive or more aggressive in nature which reflects on the life expectancy of patient. Glial cells originated brain tumors are called as Gliomas. Gliomas classified as astrocytoma and oligodendro gliomas means lowgrade gliomas such as to grade IV-high-grade glioblastoma multiform (GBM), considered as highly aggressive and primary malignant brain tumor [2]. To detect the brain tumor, magnetic resonance imaging produces wide-ranging images

---

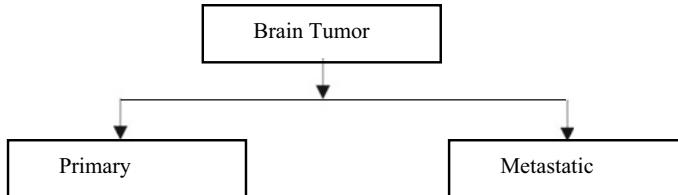
R. B. Patil (✉) · N. Ansingkar · P. D. Deshmukh  
Department of CS and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India  
e-mail: [ritapatil@gmail.com](mailto:ritapatil@gmail.com)

of the brain. Then these gliomas are treated with surgery, radiation and chemotherapy may be used as a part of treatment to slow down the growth of non-removable brain tumor. Furthermore, segmenting brain tumors from MR images has the potential to improve diagnoses, development rate calculation and treatment scheduling [3].

## 2 Brain Tumor Segmentation

It's critical to segment the tumor before applying any treatment so as to guard healthy tissues while disrupting and eliminating infected cells. Process of identifying, defining, and extrication of tumor tissues like active cells, edema and necrotic core from normal brain tissues such as CSF, White Matter, and Gray Matter is known as brain tumor segmentation [1].

Brain tumors are classified as either primary or metastatic brain tumor, depending on their originating cause [1].



1. Primary brain tumor: Brain tissue cells are considered as origin of the cells.
2. Metastatic brain tumor: In this type of brain tumor the cells get converts into cancerous at any part of body and it get spread in to brain.

## 3 Medical Imaging Techniques for Brain Tumor Diagnosis

Medical Imaging techniques like Magnetic Resonance Imaging, positron Emission Tomography, Single Photon Emission Computed Tomography, Computed Tomography provides important details about the location, size, shape of the brain tumor. These details plays vital role in diagnosis of the brain tumor in patients. From the Medical Imaging techniques the MRI is most popularly used on account of its availability and strong tissue contrast [1] but the scale of voxel value in images derived from MR images is not standardized as compared to images driven form X-ray computed tomography (CT) scans [4]. MRI considered as non-invasive in imaging technology like vivo which employs signals of radio frequency to stimulate target tissues, causing them to generate inner images in strong magnetic field. Altering the excitation and recurrence times during acquisition of image database produces images of different MRI sequences [1]. The same tumorous cells may have drastically different grayscale values according to the type of MRI device used (1.5, 3 or 7 tesla) and the technique of acquisition (view value, voxel resolution, intensity of gradient,

b0 value). Although it varies from device to device, approximately 2D images's 150 slices are formed during MRI acquisitions for reflecting the 3D brain volume. Moreover, the requisite standard modalities slices are integrated for diagnosis, the data becomes quite dense and complex.

These various MRI modalities create different forms of tissue distinction images that presents structural evidences to enable detection and segmentation of tumors in their sub region. Using many MRI modality is frequently employed for segmenting glioblastomas tumor as these are infiltrative tumors with fuzzy border and difficult to identify from healthy tissue. There are 4 MRI modalities generally used to detect Gliomas are:

1. T1 weighted Magnetic Resonance Imaging (spin-lattice relaxation) T1: T1 type of images are for characterizing healthy tissues.
2. T2 weighted Magnetic Resonance Imaging (spin-spin relaxation) T2: T2 type of images are used for define the swelling of region which generates bright signal on the image.
3. T1 weighted MRI with gadolinium contrast enhancement—T1-Gd: T1-Gd type of images are the tumor edges which can be effortlessly be distinguished with the bright signal of the accumulated contrast agent in active cell areas of the tumor tissue. Necrotic cells don't interrelate with the contrast agent, it can be analysed with hypo intense part of the tumor core to segment them from the active cell region in the same order.
4. Fluid Attenuated Inversion Recovery (FLAIR): This type of FLAIR images denotes that water molecules have suppressed and it helps in distinguishing edema region from the Cerebrospinal Fluid (CSF).

## 4 Classification of Tumor Segmentation Method

1. Manual Segmentation Method
2. Semi-automatic Segmentation Method
3. Fully Automatic Segmentation Method (Table 1).
  1. Manual Segmentation Method

The manual segmentation process of brain tumor for diagnosing cancer from MRI images obtained in clinical practice is a composite and time consuming task. The effort of MRI modalities used for glioma diagnosis presently needs manual explanation and segmentation of a huge number of various modalities MRI images in clinical practice. Detection of brain tumors at early stage can improve treatment changes and increases the patient's survival percentage. Hence, the automatic brain tumor image segmentation is required to reduce time and complexity [1].

Ali Isin proposed manual segmentation technique of brain tumor that results with BRATS dice scores is 88%, 93% and 74% for whole tumor region, core tumor region and active tumor region respectively [1].

**Table 1** Brain tumor segmentation methods classified as manual, semi automatic and fully automatic [1, 5]

S. No.	Name of the method	Procedure of segmentation	Advantages	Disadvantages
1	Manual segmentation method	In this type radiologists are required which will use multi modality imaging techniques with some physiological and anatomical data evidences. Radiologist diagnose the brain tumor by going through multiple slices if images	It is commonly used to assess results of fully automatic and semi automatic methods	Time consuming task. Manual segmentation is radiologist dependent segmentation results are subject to large intra and inter rater variability
2	Semiautomatic segmentation method	Here the method of initialization of user interaction and feedback, responses evaluation is used. Region of interest is defined to get the position of tumor. Its followed by processing automatic algorithm. Pre-processing method parameters is used to get-up the input images. Here user assesses results edit the process if not fulfilled	Less time consuming than manual methods and can obtain efficient results	Susceptible to intra and inter user variability
3	Fully automatic segmentation method	Interacting with users not obligatory. AI and previous data are joint to explain the segmentation problem	Less time consuming	Tumor sub regions are required multi modalities

However, manual segmentation is a time-consuming process, strong automatic segmentation algorithms must be developed in order to give efficient and objective results. Deep learning algorithms currently have high segmentation performance, making them good alternatives for this task.

## 2. Semi-automatic Segmentation Method

Hamamci introduced “Tumor-Cut” method of brain tumor segmentation which is semi-automatic method requires user involvement to draw the maximum diameter of tumor on input MRI images. It is a reproductive model which uses cellular automata to get the tumor’s most probable map. To get the tumor probability map, a cellular automata created seeded tumor segmentation method is run two times, first for the tumor seeds provided by the user and second is for background seeds. This method comprises running the algorithm to each MRI modality autonomously Ex-, T1, T1-Gd, T2 and FLAIR then merging the data to get the final tumor volume. It results with 72%, 57% and 59% on BRATS database dice score for whole tumor region, core tumor region and active tumor respectively [6].

Kwon et al. proposed the best performing semi-automatic generative model that generates probabilistic models using prior data like spatial extent and location of fit tissues. It performs joint segmentation and registration, in which the unknown tumor compartments are extracted using previously established atlases of healthy tissues. Converting prior knowledge into appropriate probabilistic model is a challenging task. Using this approach, the BRATS dice score shows 88%, 72% and 83% performance for tumor segmentation of whole tumor region, active tumor and core tumor region respectively [7].

Mohammad Havaei employed semi-automatic segmentation method using SVM for brain tumor segmentation in which segmentation and training implemented in the same brain with 77%, 86% and 75% of dice score performance for tumor segmentation of active tumor, whole tumor region and core tumor region respectively on BRATS dataset. It transformed segmentation problem into a classification problem and a brain tumor is segmented by training and classifying within that same brain only. For segmentation of brain tumor, machine learning classification approach requires a huge MRI database from several occurrences to train. This technique user initializes the procedure by picking a tissue type in subset of voxel from a anyone case. For subsets voxels, intensity values are extracted accompanied by spatial coordinates as features and train a support vector machine which is used to categorize all the voxels of similar image to equivalent tissue type [8].

### 3. Fully Automatic Segmentation Method

Automatic brain tumor segmentation methods of MRI can be classified as; discriminative and generative methods [1]. The discriminative method exploit previous acquaintance on the brain’s anatomy and generally relies to extraction of low-level features of image which directly creates the connection between features and the tag of a given voxel [4]. According to previous research findings the tactics based on classification techniques were very popular among other automatic approaches [5].

Majority of discriminative approaches includes steps like pre processing, Feature Extraction, Classification and post processing procedures. In the Pre processing noise removal, skull-stripping and intensity bias correction are done [9–11]. After the pre-processing Image processing methods are used to excerpt features which comprise information which is a distinguishing characteristic of an image and accurately represent each different tissue type [12].

Some examples of features include dissimilar related features, Discrete Wavelet Transforms, textons, multi fractal Brownian features of motion, statistical features, raw intensities, edge based features, local image textures and intensity gradients [13–16]. Classifier like neural networks, support vector machines, AdaBoost, and k-nearest neighbor classifier, self-organizing maps, random forests have been used by yielding acceptable results of Segmentation [12–14, 16–18].

Discriminative approaches attempt how input image and the ground truth are connected. It is based on feature selection and extraction. Mostly its rely supervised learning techniques that require huge data collection with trusted ground truth [7].

Generative models based on domain specific preceding knowledge regarding the appearance of both healthy and unhealthy means tumor tissues [4]. A typical generative model of MR brain images initiated by Prastawa et al. that record brain images onto probable map for abnormalities [19, 29]. Position of brain with a large tumor onto a template is a stimulating job, techniques perform registration and tumor segmentation simultaneously.

Tustison exploited asymmetry and statistical characteristics to train concatenated RFs by using the output of the first RF as input to the second, its considered as best successful traditional discriminative method [13].

Despite the strong performance of these standard classification approaches, a new method of segmentation techniques i.e. fully automatic brain tumor based on deep learning methods is gaining traction.

With hi-tech outcomes. The automatic Brain Tumor segmentation based on deep learning approaches have efficient processing and objective evaluation of the large amounts of MRI-based image data [1].

Convolutional Neural Networks (CNNs) is the most popular for object recognition and biological image segmentation technique [20, 21]. CNNs representative complicated features from database itself automatically. As a result of this attribute, study on CNN based brain tumor segmentation focuses e on network construction design than on image processing for extracting the features. CNNs employ image patches as inputs and extract a hierarchy of progressively sophisticated features using local subsampling & trainable convolutional filters [1].

Dvorak and Menze proposed multi-model automated brain tumor segmentation technique k-means and using local structured prediction with CNN that results 83%, 75% and 77% on BRATS dice score for whole tumor region, core tumor region and active tumor region respectively [22].

A 3D CNN architecture for the multi-modal MRI glioma segmentation task proposed by Urban in which multimodality 3D patches, cubes of voxels which is take out from various brain MRI modalities and its been inputted to CNN for forecasting tissue label of center voxel of the cube. The 4D input data is handled effectively by CNN i.e. 3 dimensional spatial intensity information and a additional dimension for MRI methods that helps to better representation of 3 dimensional type of biological structures. But it promotes load of the system to process it. They designed two different networks, one with hidden input layer contains 15 3D filters that have 53 spatial dimension and other one with secreted layer of 40 filters for 53 spatial dimension. The average result of these two networks with BRATS dice scores for

whole tumor region, core tumor region and active tumor region is 87%, 77% and 73% respectively [23].

In apposite of high dimensional method of Urban and Zikic devised an explanation method to transform 4 dimensional data such that standard 2 dimensional CNN structures may utilized to tackle the brain tumor segmentation problem [24]. It increased computational efficiency with removal of the burden of high dimensional CNN design and results with BRATS dice scores is 87.3%, 73.6% and 69% for whole tumor region, core tumor region and active tumor region respectively [25].

Davy proposed full automatic brain tumor segmentation technique using 2 pathway CNN for local and global processing that results 85%, 74% and 68% dice score on BRATS dataset for whole tumor region, core tumor region and active tumor region respectively [24].

Tustison et al. focused on fully automatic approach of brain tumor segmentation in which it concatenated RFs and accomplished using first order statistical features and asymmetry. It results with 87% 78% and 74% of dice scores for Whole tumor region, Core tumor region and Active tumor region respectively [26].

The brain tumor segmentation performance also evaluated by Sergio Pereira using deeper CNN architectures and obtained BRAts dice scores in order to whole tumor region, core tumor region and active tumor region of 88%, 83% and 77% respectively [27].

Havaei et al. implemented a cascaded two-pathway CNN architecture by extracting smaller sized patches and larger sized patches at the same time, and process local details of the brain MRI along with larger context of brain tissue. In this, CNN exploits both local features as well as more global contextual features simultaneously. It results with BRATS dice scores for whole tumor region, core tumor region and active tumor region is 88%, 79% and 73% respectively [4].

Some glioma segmentation strategies coupled the use of CNN with other classification or clustering algorithms. A local structured prediction with CNN is proposed by Pereira. The segmentation performance of this method reported using BRATS dice scores for whole tumor region, core tumor region and active tumor region is 83%, 75% and 77% respectively [27]. In some circumstances, segmentation findings are refined to improve performance. Connected components (CC) and conditional random fields (CRF) are two prevalent options [4, 15, 17, 18, 28–30].

## 5 Challenges in Automatic Segmentation

Automatic gliomas segmentation is a very tedious problem as tumorous brain MRI data is a 3 Dimensional database where tumor size, shapes and location can vary significantly from patient to patient. Tumor boundaries are frequently unclear and asymmetrical, with cutoff which offers a significant challenge to classic edge-based approaches.

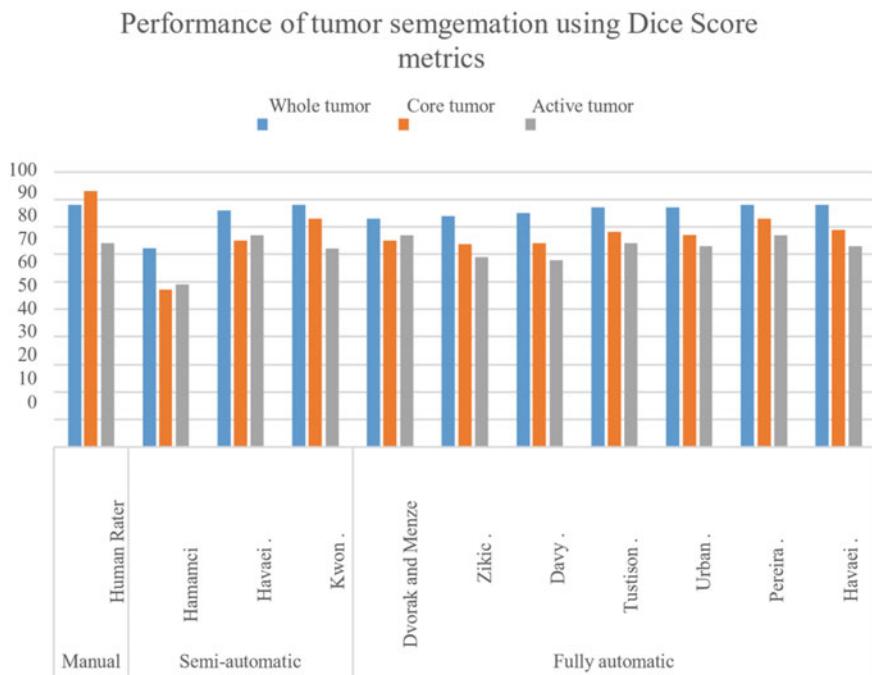
The MRI data generated from clinical scans or synthetic databases for brain tumors is inherently complicated. MRI acquisition devices and techniques can differ

considerably from scan to scan, resulting in intensity biases and other differences for each slice of image in the dataset. MRI is a practicable tool for detection of cranial sutures [31]. The need for numerous methods of Imaging to effectively segment tumor sub-regions even adds to this complexity.

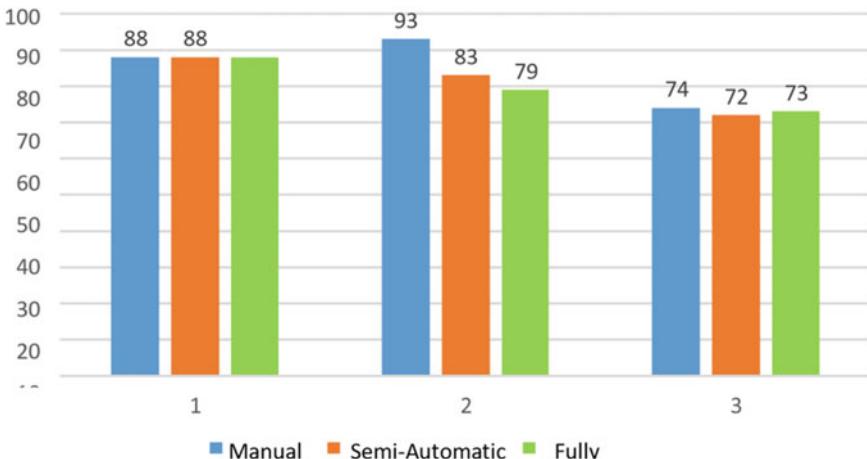
It's difficult to compare the outcomes of different brain tumor image segmentation methods to the state-of-the-art. Therefore, the performance of brain tumor segmentation techniques using dice Score performance metric for whole tumor region, core tumor region and active tumor region graphically represented in Fig. 1.

Figure 2 represents comparative study of manual, semi-automatic and fully automatic brain tumor segmentation maximum results in terms of dice score performed on BRATS dataset for whole tumor region, core tumor region and active tumor region.

Figure 2 clearly indicates that fully automatic brain tumor segmentation results better in terms of dice score as compare with manual and semi-automatic tumor segmentation method.



**Fig. 1** Performance of tumor segmentation using dice score metrics



**Fig. 2** Comparison of dice score performance metric for tumor segmentation

## 6 Conclusions

The availability of open datasets and the widely established BRATS benchmark have recently offered a common venue for researchers to build and objectively evaluate new solutions using existing techniques. This paper reviewed tumor segmentation techniques for glioma segmentation. Among the available techniques the automatic segmentation of the brain tumors for the diagnosis of cancer is extremely challenging task. Translating prior knowledge in the probabilistic maps or identifying highly demonstrative features for classifiers is a difficult challenge in old-style automatic glioma segmentation systems. The convolutional neural networks (CNN) have the benefit of learning representative complex features for both healthy and cancerous brain tissues directly from multi-modal MRI images. Results shows that fully automatic brain tumor segmentation results better in terms of dice score as compare with manual and semi-automatic tumor segmentation method.

## 7 Future Scope

This paper reviewed recent updates in the tumor segmentation methods of brain tumor for the diagnosing cancer in children. These studied methods will be taken as reference to invent a hybrid technique of segmentation to detect the brain tumor.

## References

1. A. Isin, C. Direkoglu, M. Sah, Review of MRI-based brain tumor image segmentation using deep learning methods, in *12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS*, 29–30 Aug 2016, Vienna, Austria. Procedia Computer Science, vol. 102 (2016), pp. 317–324
2. R. Stupp, F. Roila, Malignant glioma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann. Oncol.* **18**(Suppl 2), 69–70 (2007)
3. M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, PeirreMarc Jodoin, H. Larochelle, Brain tumor segmentation with Deep Neural Networks. *Med. Image Anal.* **35**, 18–31 (2017)
4. N. Gordillo, E. Montseny, P. Sobrevilla, State of the art survey on MRI brain tumor segmentation. *Magn. Reson. Imaging* **31**(8), 1426–1438 (2013)
5. Andac Hamamci, Nadir Kucuk, Kutlay Karaman, Kayihan Engin, Gozde Unal, Tumor-Cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications. *IEEE Trans Med Imaging*, Mar. 2012, 31(3), pp. 790804
6. Dongjin Kwon, Russell T. Shinohara, Hamed Akbari and Christos Davatzikos, Combining Generative Models for Multifocal Glioma Segmentation and Registration, *Medical Image Computing and Computer-Assisted Intervention— MICCAI 2014*. Springer, 2014, 17(01), pp. 763–770.
7. M. Havaei, H. Larochelle, P. Poulin, P.-M. Jodoin, Within-brain classification for brain tumor segmentation. *Int. J. Comput. Assist. Radiol. Surg.* **11**(5), 777–788 (2016)
8. R.D. Nowak, Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. Image Process.* **8**(10), 1408–1419 (1999)
9. A.H. Zhuang, D.J. Valentino, A.W. Toga, Skull-stripping Magnetic Resonance Brain Images Using a Model-based Level Set. *Neuroimage* **32**, 79–92 (2006)
10. Mohak Shah, Yiming Xiao, Nagesh Subbanna, Simon Francis, Douglas L Arnold, D Louis Collins, Tal Arbel, Evaluating intensity normalization on MRIs of human brain with multiple sclerosis”, *Medical Image Analysis*, April 2011, 15(2), pp. 267–282
11. Aparna Joshi, Dr. Vinay Chavan, Dr. Parag Kaveri, Effect of Changing Distances for Extracting Image Information for Error Reduction of Mouth Features. *International Journal of Next-Generation Computing - Special Issue*, Vol. 12, No. 2, April 2021, pp. 270–279.
12. N.J. Tustison, K.L. Shrinidhi, M. Wintermark, C.R. Durst, B.M. Kandel, J.C. Gee, M.C. Grossman, B.B. Avants, Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. *Neuroinformatics* **13**(2), 209–225 (2015)
13. V. Anitha, S. Murugavalli, Brain tumor classification using two-tier classifier with adaptive segmentation technique. *IET Comput. Vision* **10**(1), 9–17 (2016)
14. T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using threedimensional textons. *Int. J. Comput. Vision* **43**(1), 29–44 (2001)
15. Atiq Islam, Syed M. S. Reza, Khan M. Iftekharuddin, Multifractal texture estimation for detection and segmentation of brain tumors. *IEEE Trans BioMed Imaging*, Nov. 2013, 60(11), pp. 3204–3215.
16. Stefan Bauer, Lutz-P Nolte, Mauricio Reyes, Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011*, Springer 2011, pp. 354–361.
17. Darko Zikic, Ben Glocker, Ender Konukoglu, Antonio Criminisi, C. Demiralp , J. Shotton, O.M. Thomas, T. Das, R. Jena, S.J. Price, Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel, MR. *Med Image Comput Assist Interv*, 2012, 15(3), pp. 369–76
18. Prastawa M., Bullitt E., Ho, S., Gerig, G., Robust estimation for brain tumor segmentation. *Medical Image Computing and Computer-Assisted Intervention-MICCAI*, 2003, Springer, pp. 530–537.

19. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012, pp. 1097–1105.
20. Dan Ciresan, Alessandro Giusti, Luca Gambardella, Jürgen Schmidhuber, Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. *Advances in Neural Information Processing Systems*, 2012, pp. 2843–2851.
21. B. Menze, The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging* **34**(10), 1993–2024 (2015)
22. Urban G., Multi-modal Brain Tumor Segmentation Using Deep Convolutional Neural Networks. *MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS)*, 2014, pp. 31–35.
23. Zikic D., Segmentation of Brain Tumor Tissues With Convolutional Neural Networks, *MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS)*, 2014, pp. 36–39.
24. N. Tustison, Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. *Neuroinformatics* **13**(2), 209–225 (2015)
25. Davy A., Brain tumor segmentation with deep neural networks, *MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS)*, 2014, pp. 1–5.
26. S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans Med Imaging* **35**(5), 1240–1251 (2016)
27. Darko Zikic, Yani Ioannou, Antonio Criminisi, Matthew Brow, Segmentation of brain tumor tissues with convolutional neural networks *MICCAI Multimodal Brain Tumor Segmentation Challenge (BRATS)*, Oct. 2014, pp. 36–39.
28. M. Prastawa, E. Bullitt, G. Gerig, Simulation of brain tumors in MR images for evaluation of segmentation efficacy. *Med. Image Anal.* **13**(2), 297–311 (2009)
29. K. Okada, T. Soejima, H. Sakamoto, J. Hirato, J. Hara, Phase II study of reduced-dose craniospinal irradiation and combination chemotherapy for children with newly diagnosed medulloblastoma: a report from the Japanese Pediatric Brain Tumor Consortium. *Pediatr. Blood Cancer* **67**(11), e28572 (2020)
30. G. Sreedher, C. Gillespie, M. Brown, S.S. Ganapathy, Cranial suture evaluation on routine pediatric MRI. *Curr. Probl. Diagn. Radiol.* (2020)
31. C. Cahaney, P. Stefancmin, K. Coulehan, R. Parker, T. Preston, J. Goldstein, L. Hogan, T.Q. Duong, Anatomical brain MRI study of pediatric cancer survivors treated with chemotherapy: correlation with behavioral measures. *Magn. Reson. Imaging* (2020)

# Mobile Applications for Smart Agriculture in India: “An Analytical Approach”



Madhu Gaur and Preeti Tiwari

**Abstract** Agriculture in India is at a fork in the road, and the introduction of modern ICT solutions can significantly enhance crop productivity and profitability, resulting in benefits for smallholder farmers in India, where agriculture contributes for over 17% of the nation’s Gross Domestic Product. In terms of agricultural modernization, the country whose GDP is based on agriculture is incomplete. High labor costs, crop production uncertainty, lack of education and information about new methods and technology, continuing to the same traditional practices in agriculture, and inefficient irrigation system use are all factors for low productivity and crop drying may occur as a result of this unpredictability in the cultivation process. Digital India would have used the internet and mobile applications to interconnect rural Indian farmers everywhere in the world and providing them with all the resources they need to succeed in agriculture in India. The research examined at how ICT and Android applications are implemented in farming today, and the way they have revolutionized business in agriculture by bring forth a computerized platform, along with their influence in the business. This paper focuses the concept of smart farming, which utilizes mobile application technologies to support farmers with agricultural production information such as seed expenditures, temperature and humidity, composition of the soil, humidity level, weather forecast, pesticides and fertilizers utilization, knowledge of the various government schemes and facilities for farmers, and many more. This paper will also help in bringing out the current agricultural difficulties that farmers are facing. According to the survey, there are only a few agricultural based applications available in India. That provides reliable information to the farmers. This research paper focuses on the comparative analysis of various Mobile applications that supports the farmers in their various agricultural activities. Number of features the applications like their precision level, data reliability, certificated contents availability and also the services for specific geographical locations is compared. So as to propose an application with better features and better farming support. There are

---

M. Gaur (✉) · P. Tiwari  
International School of Informatics and Management, Jaipur, India  
e-mail: [gaurmadhu91@gmail.com](mailto:gaurmadhu91@gmail.com)

P. Tiwari  
e-mail: [preeti.tiwari@icfia.org](mailto:preeti.tiwari@icfia.org)

a number of mobile applications available to improve the current state of farming; a summary of all applications is provided below.

**Keywords** Agriculture in India · Smart farming technologies · Mobile applications · Information

## 1 Introduction

India is a large country, and agriculture is one of the most important markets in our countryside, it significantly defines country's economic and social development. In regards of farming, India has been ranked second in the world with respect to farming itself. Agriculture is the main source of income of the country, with approximately 58% of the people depending on it directly, according to the most recent survey [1]. Technology has made communication easier and more efficient, and it is also playing an important part in agriculture also. The traditional farming techniques have changed as technology has infiltrated the agricultural industry as well. There are high-tech approaches to locate the greatest quality seeds and latest farming techniques but there seems to be a lot that can still be done with the advancement of technology. The upcoming technologies in agriculture can be used to predict that how to enhance farmer profitability [1]. While there are a variety of mobile applications available to help improve the current state of agriculture in India, where Farmers in developing countries keep practicing old methods of agriculture due to a lack of education and awareness of current technologies. A number of technologies have been developed to make farming digital. By utilizing these technologies, farmers can improve their abilities to do farming. This study examines the developments in various technological fields related to smart farming. There are a few of them that are related to smart farming applications. An android application is mainly used for sustainable development of farmers' businesses. The use of these applications is very beneficial for avoiding such kind of problems because a lot of the time farmers are confused as to which fertilizer, pesticides and when to apply particular farming actions will be most beneficial for them [11]. Using the digital farmer application, we can find information about different crops and other agriculture products for farmers. It is an android-based application that provides information to farmers [12]. To build and analyse the needs of farmers can be fulfilled by android applications that contain information about crops, fertilizers currently being used in market, providing notifications for the incremental growth of crops. The objective of this study is to examine the use and limitations of mobile applications in the agricultural sector, where the global mobile ecosystem exists across a variety of platforms, and the problems associated with scaling up mobile agriculture applications. It also aims at finding out how Smartphone agriculture applications are doing right now. It describes the characteristics of several apps currently in use. Throughout this research, the most popular apps are factored into the equation.

**Table 1** Comparison between smart agriculture and traditional farming

Agricultural requirements	Traditional farming	Smart farming
Time	More time	Less time
Yields	Low	High
Farming cost	High cost	Low cost
Data acquisition	Very low	Comparatively high
Monitoring of field	Very high	Low monitoring
Utilization of water	High	Low
Staff needed	Yes	No
Production sustainability	High	Low
Profitability	Low	Medium to high
Economic sustainability	Low	High

## 2 Comparison of Traditional Farming and Smart Farming

There are common methods followed by farmers for the cultivation of crops that cause lack of utilization of effective resources like man power and water recourses. There are numerous different strategies are utilized by farmers today. Agriculture machine accompanied with the aid of using conventional strategies effects in excessive investments value and additionally artificial destruction [13]. Whereas Smart farming is the alternate technique for the utilization of water in an effective to cultivate a crop. Smart farming helped to reduce the limitations of traditional farming by using real time monitoring, time, accurate information, how can a farmer protect the crop from diseases. Complete knowledge of pesticides and other useful information can easily through smart android apps. Table 1 shows the key difference of Traditional farming and Smart farming.

## 3 Role of Mobile Applications with Reference to Farming in India

Indian Agriculture is progressively becoming much more technologically sophisticated. The Digital India platform, which was established in 2015 by Indian Prime Minister Narendra Modi, focused on promoting technological literacy and the development of digital infrastructure in India, with the purpose of empowering people in rural areas so that community of farming can develop. Since, agriculture remains an important source of income for 58% of Indian households. The role of mobile

apps in the agriculture sector will significantly help in increasing the productivity and sustainability.

Agriculture applications, on the other hand, are the most convenient and useful medium for educating and guiding farmers on how to use technologies for better farming [15]. It provides a comprehensive reference to scientific farming, including crop development, harvesting of any crop, plant protection, understanding of current trends and practices of agriculture. The mobile apps can enable farmers to handle their farming challenges effortlessly. These challenges may arise because of variety of reasons like pest attack, diseases or insect infection, productivity yield to name a few. An agriculture app might be a farmer's best friend, allowing them to increase the output of their field without investing any money. These apps can be used and downloaded from the Google Play Store without having to spend anything [5].

There are a number of mobile applications that are employed all over the world for various purposes, including farming, but their use in the United States is still constrained [13]. In this paper it is elaborated various agriculture mobile application that are potentially can be used in farming. These agriculture applications are extremely significant for Indian farmers and the agriculture community in our country because they keep them up to date on the latest technology and trends in agriculture sector that has become extremely important for everyone. Smartphone apps also help to overcome the educational and entertainment divide between rural residents and the government. These apps are also suitable for obtaining up-to-date market rates, government policies and schemes for farmers, weather-related information, the latest technologies and films for learning, agriculture-related news, and so on [4]. By using these apps, they can easily start to ask questions from agriculture professionals, experts and also can communicate with live videos in the order to acquire knowledge.

## 4 A Survey on Mobile Agricultural Apps Used in India

A research was conducted to assess the characteristics of Android agricultural production apps most of which are found on well-known app stores. This mostly focused on applications accessible in the App Store, Google Play Store, and Windows Phone Store [17]. The results show that there are as in almost 551 Android apps and 589 iOS apps in various categories, which would include animal production, financial focused primarily, farm management crops, agriculture technology and innovations, agriculture machinery, weather reporting Apps, training-related apps, agricultural media coverage, and other challenges [6].

## 5 Architectures Supporting Android

Software stack of components of android architecture designed to meet the needs of mobile devices. It includes a Linux kernel, as well as C/C++ frameworks and libraries through application framework services, runtime, and applications [7]. In the form of Java classes, the application framework layer includes numerous higher level services to the application. Activity Manager, Content Provider, Resource Manager, Notification Manager, and View System are just a few examples. As a developer, Android is an open source platform that allows to create extraordinarily powerful applications. It allows the developer to take full advantage of hardware, access location data, and execute background services, among many other things. Once the application is published, it can be effortlessly downloaded by any third party through online stores such as Android stores and Google's app store.

### 5.1 *Agri App*

Farmers could use the Agri App to know about the current agriculture trends. It helps farmers to understand about the methodologies involved in soil selection and understanding, as well as environmental parameters for various crops. In 2017, Agri App seems to have been a finalist for the Digital Empowerment Foundation's Billionth Award in the field of agricultural and environmental [7]. Agri App Technologies Pvt. Ltd developed this Android-based app, which would also be effective for controlling plants against diseases, weeds, and other pests. Package services contain information about 83 different crops. Chat with experts allows people to communicate directly with crop expert scientists via text and phone calls. This supports the user in understanding the evolution of agriculture and new technologies [7].

### 5.2 *IFFCO Kisan*

Indian Farmers Fertilizer Cooperative Limited is known as Kisan (IFFCO). This software improves the quality of life of community by providing a wide range of possibilities [12]. The software is specifically designed for farmers; therefore it has a number of important functions for farmers. IFFCO Kisan Agriculture App is a mobile-based service in India that promises to support Indian farmers in factors includes the following decision making by providing them with highly personalized relevant information to their requirements. Farmers could get regular updates and suggestions for their day-to-day farms and agriculture challenges using this app [8].

### **5.3 Agri Media**

Agri Media is a video app dedicated specifically to Indian farmers. The application was created for the benefit of Agricultural production by Digital Agri Media. Farmers might just get updates in the form of videos, and it's very easy to use. This is one of the greatest agriculture apps manufactured in India. All of the videos in the app are available in Hindi and Gujarati, and will be available in different regional languages in the coming years to assist farmers in better farming. The objective of the application is to assist farmers in strengthening their farming skills and boosting their revenues [9].

### **5.4 Farm Bee**

Farm Bee is a free app that can be downloaded from the Google Play Store. This application is primarily for enhance the effectiveness of pomegranates by better managing the farm utilizing Farm Bee. The app includes important information regarding pomegranate farming, including weather conditions, soil moisture levels, and atmospheric conditions, all adapted to the requirements of farmers. Sweetchillie Technologies is the firm behind Farm Bee is the Pomegranates Expert. 6.1.0 Was the latest edition released by the developers.

### **5.5 Kisan Yojana**

Kisan Yojana is a multifunction mobile app developed by the Ministry of Agriculture and Farmers Welfare's Department of Agriculture and Cooperation to support farmers by providing information. On the kisan app, customers may get a regular weather report as well as a five-day before weather prediction for the specified location. Dealers of seeds, insecticides, fertilizers, and farm machinery are also included in their names, phone numbers, and locations. This application is also available on the Google Play Store for Android users. And all of the characteristics are crucial in every manner.

### **5.6 Agri Market**

Farmers would be using the Agri market mobile app to see just what crops are selling in markets within 50 miles of geographical location. This application utilizes mobile GPS to automatically capture a person's precise location and retrieve all of the crop market price in the nearest market. The goal of this app is to keep farmers

knowledgeable about crop prices in their area. The application is supported in English and Hindi languages [8].

### **5.7 *Kheti Badi***

This application, on the other hand, is the only one of its sort dedicated to farmers. The fundamental problem for our country is the information and knowledge gap between industrialized and developing countries. Agriculture traders and farmers in India will get benefits from the application. They can now get internet prices by using this application. Almost all agricultural commodities from all around the Indian market are covered by this app. Farmers can use this application in the order to transform the chemical farming in to organic farming making the livelihood better, and will connect to the customers in order to make money [9].

### **5.8 *Kisan Suvidha***

Prime Minister Narendra Modi launched the Kisan Suvidha app in 2016 in order to work more toward the empowerment of farmers and the development of communities. The App's layout is straightforward and user-friendly, and it gives facts concerning the modern temperatures in addition to climate predictions for the following numerous days, marketplace charges within side the close-by community, fertilizers, seed, and equipment expertise, and so on [13].

### **5.9 *Crop Insurance***

It's fantastic software that assists farmers in calculating insurance premiums for alerted crops as well as providing information on cut-off dates and company contact information for their crop and location. Farmers can use that as a reminder and calculation for their insurance. It may be used to attain the data on any notified area's traditional sum insured, prolonged sum insured, top rate details, and subsidy data [18].

## 6 Comparative Study of Android-Based Apps for Farmers in India

In India, there seem to be a variety of agriculture-related mobile apps available. These apps can be obtained from the internet. Some apps are pre-installed on Smart phones, while others must be downloaded from an application distribution platform. The table shows the features of different app used in India. Table 2 shows the results. A comprehensive comparison of agriculture-related Android apps.

## 7 Drawback in above Mobile Applications and purpose of this study

Existing mobile applications generally have their own distinguishing qualities. Each of the apps seems to have its own different advantages and disadvantages in contest to farmers. Two-way communication is not provided by any of the developers, according to the study of numerous farming applications. However there is no Android app that supports toll-free numbers. As a result due to the excessive charges, farmers find it difficult to make calls and conclusions. Because not all the farmers are good in English and some are also not even in Hindi, the apps should have been in the farmers' native language, or there should be an option to change the language according to their requirements. As of today, none of the Android Applications available for this study provide farmers with the ability to purchase and sell agricultural products via digital marketing. An application that is user-friendly and can fulfill all requirements on a single platform should be developed [11].

## 8 Conclusion

Different varieties of agriculture mobile phone applications have been thoroughly discussed, including information-related apps, weather apps, NEWS apps, and E-Government apps. The greatest agriculture app obtainable in India right now, thus according government scheme, is Kisan Yojana. Whereas Agri App fits best App in user's point of view. There are different applications have been developed and have been used for multiple purposes. According to their functionalities, each of these apps has a particular use. A variety of apps are being used for many functions related to farmer activities, such as crop information, fertilizer, seeds, crop selling, market price, irrigation information, crop production estimation, weather information, and information on optimum agricultural techniques. Many of the apps are discovered to be static in nature, meaning they don't execute any procedures and therefore are simply used as information assets. Instead, dynamic apps will become more convenient to use. Also, if all of the above-mentioned features are combined into a

**Table 2** Comparison of available apps for Indian agriculture

S. No.	Android app name	Feed back form	Text	Audio and video	Image	Significant features
1	Agri App	Yes	Yes	Yes	Yes	Download information, place order, group chat, make calls, IVRS and other
2	IFFCO Kisan app	Yes	Yes	Yes	Yes	Latest mandi price, best practices tips
3	Agri Media	Yes	Yes	Yes	Yes	Videos, news, farmer's question expert review
4	Farm Bee	–	Yes	–	–	For pomegranates farms, informational
5	Kisan Yojana	Yes	Yes	Yes	Yes	Government schemes, market prices etc.
6	Agri Market	–	Yes	–	Yes	Market price of crops within location 50 km
7	Kheti Badi	–	Yes	–	–	Info related organic farming, government schemes, real time price in market
8	myRML	Yes	Yes	–	Yes	Group chat and GPS location, information
9	Modern Kheti	–	–	–	–	Can download information about seeds
10	Water Reporter	Yes	Yes	–	Yes	GPS location
11	Crop Info	–	–	–	–	Downloading information
12	E- Agro	Yes	Yes	Yes	Yes	Email, Make calls, IVRS, Video calling

(continued)

**Table 2** (continued)

S. No.	Android app name	Feed back form	Text	Audio and video	Image	Significant features
13	Kisan Books	Yes	Yes	–	Yes	Place order of material, make call to vendors
14	Fertilizer calculator	–	Yes	–	–	Information about fertilizer, GPS
15	My Soil	–	–	–	–	GPS enable, humidity level
16	GAF 14	Yes	Yes	–	Yes	Group chat, GPS location
17	Farm emergency contacts	–	Yes	–	–	Make calls to nearest vendors
18	Farm Progress	–	–	–	–	Download information
19	Agriculture policy research	–	–	–	–	–
20	Crop and medicinal plants	–	–	–	–	Downloading information, pesticide information

single app, and if these apps are available in farmers' native languages as we know the literacy rate in India is low in contest to farmers, particularly in rural areas. It will be simpler to be using mobile apps when they are in their own language. Mobile Apps have such a number of drawbacks, including language dependencies, static information, and a lack of participation in online marketing. They have been discovered to be the best source for agricultural intensification.

## References

1. M. Kaur, O.P. Sharma, G. Mohan, Mobile apps: role in Indian agriculture. *Biotica Res. Today* **2**(8), 797–799 (2020)
2. S.A. Kumar, C. Karthikeyan, Status of mobile agricultural apps in the global mobile ecosystem. *Int. J. Educ. Dev. Using Inf. Commun. Technol.* **15**(3), 63–74 (2019)
3. A.K. Singh, R. Singh, P. Adhiguru, R. Hazare, *Krishi Gyan Mobile Apps* (2018)
4. A. Barh, M. Balakrishnan, Smart phone applications: role in agri-information dissemination. *Agric. Rev.* **39**(1), 82–85 (2018)
5. A. Balkrishna, J. Sharma, H. Sharma, S. Mishra, S. Singh, S. Verma, V. Arya, Agricultural mobile apps used in India: current status and gap analysis. *Agric. Sci. Dig.* **41**(1) (2021)
6. J. Akinode, S. Oloruntoba, An Android Based Farm Products Information System and Extension Services

7. S. Kaur, K.S. Dhindsa, Comparative study of android-based M-Apps for farmers, in *International Conference on Intelligent Computing and Applications* (Springer, Singapore, 2018), pp. 173–183
8. S.K. Mahapatra, Smartphone apps for agri-information dissemination during Covid19 lock-down. *Biotica Res. Today* **2**(5), 116–119 (2020)
9. S.A. Kumar, C. Karthikeyan, Eminence of mobile agricultural apps in the worldwide mobile network. *Int. J. Dev. Extens.* **9**(1) (2020)
10. T. Popović, V. Maraš, S. Čakić, S. Šandi, S. Radonjić, K. Pavićević, Use of mobile applications in smart agriculture, *Practical Guide for the Use of ICT in AET*, p. 36
11. L. Raja, S. Vyas, The study of technological development in the field of smart farming, in *Smart Farming Technologies for Sustainable Agricultural Development* (IGI Global, 2019), pp. 1–24
12. N. Monali, K. Snehal, N. Mayur, G.B. Chavan, Evolution from traditional farming to digital farming. *Evolution* **5**(03) (2018)
13. R.N. Athirah, C.Y.N. Norasma, M.R. Ismail, Development of an android application for smart farming in crop management, in *IOP Conference Series: Earth and Environmental Science*, vol. 540, No. 1 (IOP Publishing, 2020), p. 012074.
14. C. Gupta, M. Gupta, P. Joshi, A. Kumar, Information and communication technology in agribusiness: a study of mobile applications in perspective of India. *J. Appl. Nat. Sci.* **13**(2), 766–774 (2021)
15. H. Patel, D. Patel, Survey of android apps for agriculture sector. *Int. J. Inf. Sci. Technol.* **6**(1–2), 61–67 (2016)
16. S.G. Karkhile, S.G. Ghuge, A modern farming techniques using android application. *Int. J. Innov. Res. Sci. Eng. Technol.* **4**(10), 10499–10506 (2015)
17. M. Singh, K. Verma, A. Shukla, Krishi Ville—Android based solution for Indian agriculture, in *2011 Fifth IEEE international conference on advanced telecommunication systems and networks (ANTS)* (IEEE, 2011), pp. 1–5
18. M.S. Farooq, S. Riaz, A. Abid, K. Abid, M.A. Naeem, A survey on the role of IoT in agriculture for the implementation of smart farming. *IEEE Access* **7**, 156237–156271 (2019)

# Offline Handwritten Character Recognition of Gujarati Characters Using Convolutional Neural Network



Pinal J. Patel, S. Kheni Rukshmani, Urmila Patel, Darshil P. Patel,  
Kinjal N. Patel, and Kinjal A. Patel

**Abstract** Handwritten character recognition (HCR) system recognizes characters in handwritten documents. There are two types of the HCR—Online handwritten character recognition, which uses data from real-time machines, and offline handwritten character recognition, which uses data from scanned documents or photographs. As different people have different ways of writing, offline handwriting recognition is difficult in comparison. A number of studies have already been conducted in different languages such as Hindi, Kannada, Telugu, Tamil, Bangla etc. However, there is still an insufficient quantity of academic work done on Gujarati script, which is the primary purpose of this paper. In this study, we present a novel image character dataset in Gujarati, which has 60 images per character. Handwritten Gujarati characters are recognized using the Convolutional Neural Network (CNN). The accuracy of recognizing handwritten Gujarati characters is greater than 90%.

**Keywords** Image processing · Deep learning · Convolutional neural network · Otsu's thresholding method · Optical character recognition

---

P. J. Patel (✉) · S. Kheni Rukshmani

Computer Engineering Department, Government Engineering College Gandhinagar,  
Gandhinagar, Gujarat, India

e-mail: [pinalpatel@gecg28.ac.in](mailto:pinalpatel@gecg28.ac.in)

U. Patel

Electronic and Communication Department, U.V. Patel College of Engineering, Ganpat  
University, Gozaria, Gujarat, India  
e-mail: [urmila.patel@ganpatuniversty.ac.in](mailto:urmila.patel@ganpatuniversty.ac.in)

D. P. Patel

Information Technology Department, Gujarat Technological University, Ahmedabad, Gujarat,  
India

K. N. Patel

Information Technology Department, Parul University, Vadodara, Gujarat, India

K. A. Patel

Computer Applications and IT Department, Gujarat Law Society University, Ahmedabad,  
Gujarat, India

## 1 Introduction

In the 1950s, work on a handwriting recognition system began. Since the previous few decades, computers have become more proficient at interacting with humans and the natural world, such as speech recognition, handwriting recognition, gesture recognition, and so on. Characters are written by humans and then recognized by machines in handwritten character recognition. Because handwriting changes from person to person, research in this area is difficult.

In India, the people of Gujarat use the Gujarati language in their day-to-day life. There is a rich literary heritage available in Gujarati script. However, due to a huge number of visually identical characters, a higher number of classes, and touching and broken characters, research in the subject of Gujarati script identification has significant challenges [1]. There is no standard dataset for Gujarati script that can be used for testing and experimentation. The recognition of Gujarati script has received very little attention. Gujarati contains thirteen vowels and thirty-four consonants, as indicated in Figs. 1 and 2. Other symbols called vowel modifiers are also used to show the connection of vowels to core consonants. Because several Gujarati characters have extremely similar appearances, they can be readily misclassified if there is enough noise [2].

The method for recognizing consonants in Gujarati script is described in this paper. The dataset contains 2040 images of 34 different classes. The sample of collected dataset is shown in Fig. 3. The suggested approach uses Deep learning architecture to identify the characters. Deep neural networks do not require explicit feature descriptions; instead, they build the best features from raw pixel data and then use those features to classify the input into multiple groups. Convolutional Neural Networks are a type of deep neural network that works with fewer parameters and is easier to train [4].

**Fig. 1** Gujarati consonants [3]

Gujarati Consonant												
ક	ખ	ગ	ધ	ચ	છ	જ	ઝ	ઝા	ટ	ઠ	ડ	ઢ
ka	kha	ga	gha	cha	chha	ja	za	zha	ta	tha	da	dha
ણ	ત	થ	દ	ધ	ન	પ	ફ	બ	બ્ર	મ	મ્ર	ય
aNa	ta	tha	da	dha	na	pa	fa	ba	bha	ma	ma	Ya
ર	શ	વ	સ	શા	ષ	હ	ષા	ળ	શ્લ	શી		
ra	la	va	sa	sha	shha	ha	ala	ksha	gna			

**Fig. 2** Gujarati vowels [3]

Gujarati Vowel							
એ	આ(ા)	ઇ (િ)	ઓ(ો)	ઓ(ુ)	ઉ (ુ)	ଓ (૦)	રિ
a	aa	e	ee	u	oo		
એ(ે)	એ(ો)	ઓ(ો)	ઓ(ુ)	ઉ	ଓ(૦)		
a	ai	o	au	am	ah		

**Fig. 3** Sample of collected dataset

સ	સ	સ	સ	સ
અ	અ	અ	અ	અ
ઝ	ઝ	ઝ	ઝ	ઝ
એ	એ	એ	એ	એ
રી	રી	રી	રી	રી

## 2 Literature Survey

Joshi and Risodkar [5] uses Nearest Neighbour classifier to classify the handwritten Gujarati characters. Filtering, edge detection, and morphological transformation have been used as feature extraction approaches. And achieved 78.6% accuracy. Bharat et al. [6] have collected dataset of 170 images of Gujarati handwritten characters of 34 different classes. They have achieved 99.4% accuracy by taking 22 features of GLCM (Grey Level Cooccurrence Matrix) when it is used in conjunction with DTW (Dynamic Time Warping). Patel and Desai [7] used combination of statistical and structural features for classification and recognition of Gujarati characters. They have used KNN classifier and achieved accuracy of 63%.

Sharma et al. [8] used three different types of features and their fusion to distinguish isolated Gujarati handwritten characters. They compared the outcomes of Support Vector Machine (SVM) and Nave Bayes (NB) classifiers for classification. Prasad and Kulkarni [9] used Fuzzy Hedges to build an Adaptive Neuro Fuzzy Classifier (ANFC) for Gujarati character recognition (FHs). They have achieved 58.78% of accuracy using ANFC-FH and 68.67% accuracy using feature selection with ANFC. Macwan and Vyas [10] used different feature extraction techniques (Transform domain, Geometric method, Structural method and Statistical method) and compared the results. SVM classifier is used to classify the handwritten Gujarati characters.

### 3 Proposed System

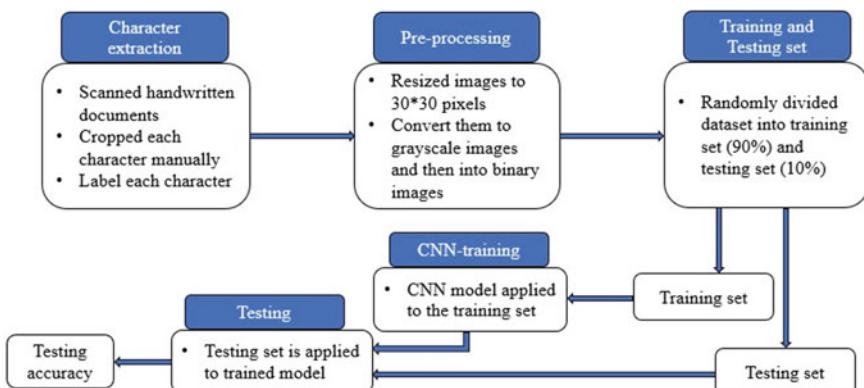
Figure 4 depicts the system's complete flow diagram. The dataset preparation phase includes character extraction, preprocessing, and the separation of the training and testing sets. CNN is used to recognize characters from the prepared dataset.

#### (1) Dataset Preparation

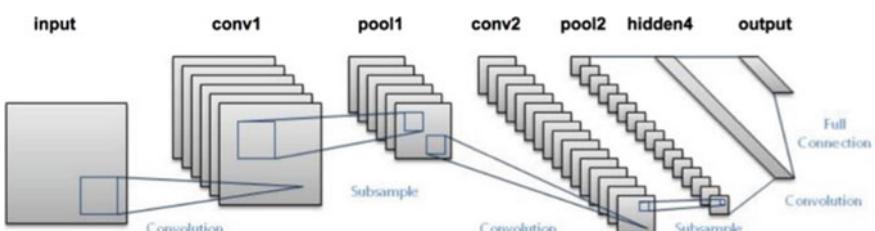
First step of dataset preparation phase is character extraction in which the handwritten documents are scanned and each character is cropped manually. After that labeling of characters is done. Dataset contains unique images. In the preprocessing step all images are resized into  $30 * 30$  pixel. Images are converted into the grayscale images and then into the binary images using Otsu's thresholding method. The prepared dataset is randomly divided into training (70%) and testing sets (30%).

#### (2) Character Recognition

In our system, the handwritten Gujarati character is identified using a deep neural network—a convolutional neural network. The simple architecture of CNN similar to our recognition system is as shown in Fig. 5.



**Fig. 4** Flow diagram of character classification



**Fig. 5** A simple CNN architecture

In general, CNNs are made up of numerous convolution and subsampling layers between the input and output layers for image processing. Following these layers are completely connected layers, resulting in a separate representation of the input data.

#### (A) Convolution Layer

Our model's most fundamental layer is the convolution layer. Convolution layers' main objective is to extract features from the input. Learnable convolution kernels and convolution operations are used to achieve this goal. When input is provided to a convolution layer, all of the kernels are convoluted with the input, and one output is generated for each kernel [11]. The first convolution layer receives the input image.

#### (B) Subsampling Layer

In CNN, it comes just after a convolution layer. The major function of the subsampling/pooling layer is to reduce the number of parameters that the network must learn.

#### (C) Fully-connected Layer

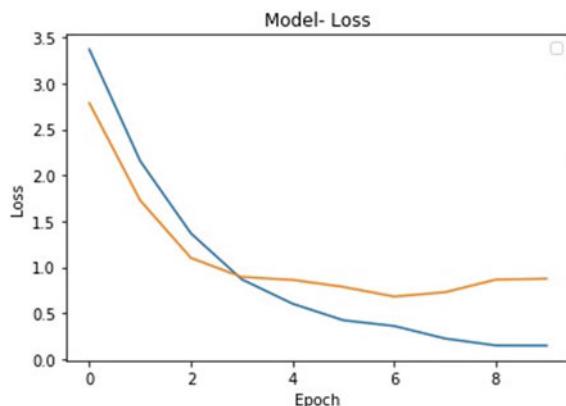
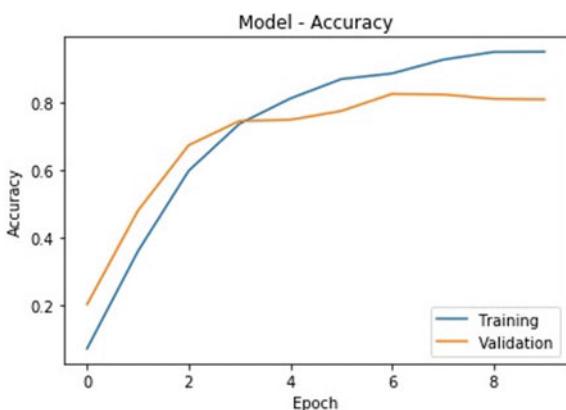
The final layers of CNN are fully connected layers. The output of the last pooling or convolution layer, which is flattened and then fed into the fully connected layer, is the input of the fully connected layer. The final Fully-connected layer provides the output of our recognition algorithm.

## 4 Results

The dataset used in this research is of 2040 (34 \* 60) images. The dataset is collected from different writers having different writing styles, different thickness of characters and of different sizes. After splitting the dataset randomly into training and testing set, Convolutional Neural Network is applied on the dataset and model gives accuracy of 82%. Model loss and model accuracy is shown by Figs. 6 and 7. This research recognizes only Gujarati consonants.

## 5 Conclusion

In this paper, the problem of handwritten character recognition for Gujarati characters is addressed. We created the dataset of 2040 images of 34 classes for Gujarati consonants. Dataset consist of many characters which are visually similar so it is difficult to identify. We used Deep learning method Convolutional Neural Network to recognize the character and we have achieved more than 90% accuracy for the characters listed in Table 1.

**Fig. 6** Accuracy of model**Fig. 7** Loss in model**Table 1** Results of model

Sr. No.	Character set	Accuracy (%)
1	શ, ષ, ગી, ધ, પી, ટી	95.83
2	ચ, થ, પ, મુ, પુ	92.59
3	દ, ણ, બુ	93.33
4	સ to શી	92.16

## 6 Future Work

The dataset of this research contains only 34 Gujarati consonants. So, it will be necessary to expand the dataset in the future which includes more images of handwritten Gujarati characters and also handwritten Gujarati vowels. One direction for future work is to focus on working with the modifiers along with Gujarati characters.

## References

1. V.L. Sahu, B. Kubde, Offline handwritten character recognition techniques using neural network: a review. *Int. J. Sci. Res. (IJSR) India* **2**(1) (2013). Online ISSN: 2319-7064
2. J.R. Prasad, Image normalization and preprocessing for Gujarati character recognition. *IJCSN Int. J. Comput. Sci. Netw.* **3**(5) (2014)
3. A.A. Jain, H.A. Arolkar, A survey of Gujarati handwritten character recognition techniques. *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)* **6**(IX) (2018)
4. S. Acharya, A.K. Pant, P.K. Gyawali, Deep learning based large scale handwritten Devanagari character recognition, in *9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, IEEE, 2015
5. D.S. Joshi, Y.R. Risodkar, Deep learning based Gujarati handwritten character recognition, in *International Conference on Advances in Communication and Computing Technology (ICACCT)*, IEEE, 2018
6. S.B. Sunilkumar, R. Yadav Arvind, M. Patel Mihir, M. Patel Mitul, Gujarati handwritten character recognition using grey level co-occurrence matrix and dynamic time warping technique, in *International Conference on Smart Technologies For Smart Nation (SmartTechCon)*, IEEE, 2017
7. C. Patel, A. Desai, Gujarati handwritten character recognition using hybrid method based on binary tree-classifier and K-nearest neighbor. *Int. J. Eng. Res. Technol. (IJERT)* **2**(6) (2013) . ISSN: 2278-0181
8. A. Sharma, P. Thakkar, D.M. Adhyaru, T.H. Zaveri, Features fusion based approach for handwritten Gujarati character recognition. *Nirma Univ. J. Eng. Technol.* **5**(2) (2016)
9. J.R. Prasad, U. Kulkarni, Gujarati character recognition using adaptive neuro fuzzy classifier with fuzzy hedges. *Int. J. Mach. Learn. Cybern.* **6**(5), 763–775 (2015)
10. S.J. Macwan, A.N. Vyas, Classification of offline Gujarati handwritten characters, in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2015
11. M.B. Mendapara, M.M. Goswami, Stroke identification in Gujarati text using directional feature, in *International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)*, IEEE, 2014
12. G. Lin, W. Shen, Research on convolutional neural network based on improved Relu piecewise activation function, in *8th International Congress of Information and Communication Technology (ICICT)*. Procedia Computer Science, vol. 131 (2018), pp. 977–984
13. S.S. Rosyda, T.W. Purboyo, A review of various handwriting recognition methods. *Int. J. Appl. Eng. Res.* **13**(2), 1155–1164 (2018). ISSN 0973-4562
14. K. Peymani, M. Soryani, From machine generated to handwritten character recognition: a deep learning approach, in *3rd International Conference on Pattern Recognition and Image Analysis (IPRIA 2017)*, IEEE, 19–20 April 2017

# Forecasting Solar Energy on Time Frame: A Review



Ashok S. Sangle and Prapti D. Deshmukh

**Abstract** For installation of solar photovoltaic system there will be adequate approaches required those mainly including solar radiation mapping, forecast, site evaluation, and potential assessment, to resolve such approaches there will be problem solving advanced techniques being evolved those are based on geospatial and machine learning technology. A review to make up the clarity related to the domain a specific role of these technologies as problem solving competency in relation to photovoltaic solar power systems (PV System), forecast and for the conversion of solar energy into electricity scenario. The review was performed by classifying previous Geospatial and Machine learning (ML) based studies according to the complexity of the active Geospatial and Machine Learning based techniques being utilized at extent to support for this stuff that is the data source, and the findings of the study area potential for more accurate assessment is required. The Geospatial technology is appropriate for handling location based and Machine Learning is more suitable for forecast related stuff but data related to solar resource and site suitability circumstances on various scales required. The claims of Geospatial Technology with Machine Learning based approaches in solar power system installation, planning and potential measure analysis of specific convinced technology, its role and truthfulness can be extended further.

**Keywords** Solar radiation · Irradiation · Forecasting · Machine learning · Geospatial technology · Photovoltaic system · Smart city · Photovoltaic cell

---

A. S. Sangle (✉)

Department of Computer Science and IT, Dr. B. A. Marathwada University, Aurangabad, India  
e-mail: [ashoka4it@gmail.com](mailto:ashoka4it@gmail.com)

P. D. Deshmukh

MGM's Dr. Pathrikar College of Computer Science and IT, Aurangabad, India

## 1 Introduction

The amounts of solar radiation are falling in numerous regions of the state/area. A right forecast gives right moves. Solar forecast is influenced by the movement of planet and meteorological/weather/temperature conditions of climatology. With the use of Machine Learning diverse weather conditions (wind, Cloudy, clear sky, sun, rain) and influence of atmospheric conditions on solar radiation can be taken into considerations. Forecasting appears on the basis weather/climatology factors, but the challenge is that not certain atmosphere constant always, over time it continuously changes so its challenge to analyses its accuracy stuff. To attain this edge-cutting role of analysis and accuracy a Geospatial and Machine Learning based approaches have been connected viably for sound sun based control framework for the concern of arranging and planning solar associated resources.

This review is being reviewed in view to explore the innovative ways to utilize or measure the sun light hitting rooftop extent for solar energy potential assessment and planning of photovoltaic (PV) system installation practices. To investigate imaginative ways to utilize sunlight hitting rooftop extent for solar energy potential implies the establishment of photovoltaic (PV) units. The forecast of solar energy will contribute for solar power productions stuff so now discovery of entire rooftop area of institutional, social and financial infrastructure of urban (smart city) with the assistance of Geospatial Technology, so it will contribute in calculating the size of the structures/unit. The average solar radiation over a household location can offer assistance to predict the aggregate electricity will be generated by solar panels and an accurate solar forecast can offer its potential, assist smooth operations and assistance to harvest the solar energy suitably, further calculation of unit size promote unit level potential, return on investment and framework load quantifications etc. It will moreover give essential information to utilities such as like electricity grids to assist its operations smoothly. We are progressing to learn more almost the current capabilities to forecast accurately the variable outputs of solar systems over an hour ahead and days ahead on time frame or horizons outline.

### 1.1 *Objective*

The key objective of review paper is to find out Data type/resources, Techniques and findings to explore the solar energy potential assessment and forecasting.

The content of this review paper organized in to five sections namely Abstract, Introduction, Literature Review, conclusions followed by references.

## 2 Literature Review

During the literature reviews following Data type/sources, techniques and findings are observed useful for achieving research objective.

### 2.1 *Data Type/Sources*

Many researchers have used Satellite image, DEM, DSM and radiation data Solar Map [1], IRS-5 CARTOSAT-1 stereo images [2]. Satellite Data [3] multispectral spatial resolution imagery [4], National Solar Radiation Database (NSRD) National Climate Data Center (NCDC). National Oceanic and Atmospheric Administration (NOAA) [5, 6]. Google EarthTM-IKONOS and GE-1 Specifically from GeoEye's GE-1 satellite and weather data [7]. LiDAR Satellite Data [8, 9], and Special Sensor and Microwave/Imager SSM/I data [10].

### 2.2 *Techniques and Findings*

Yosoon Choi, Jangwon Suh and Sung-Min Kim were utilized methods for site evaluation i.e. Boolean overlay, Weighted sum, Analytic Hierarchy Process, Fuzzy and for Radiation, Potential Assessment artificial neural network (ANN), Shadow analysis, Model validation are used [1].

Kakoli Saha By Image preprocessing technique automated the extraction of urban roof area [2].

S. Sanders, C. Barrick, F. Maier and K. Rasheed used Machine learning in Weka shown here one-hour and 24-h predictions and indicates that the Random Forest ML model yielded with improvement [5].

Xiaoyan Shao, Siyuan Lu, and Hendrik F. Hamann Forecast based on Solar Irradiance use of multi-model blending approach. Combing multiple meteorological models including numerous atmospheric and cloud prediction models based on satellite imagery as well as numerical weather prediction (NWP) were utilized to explore the Forecast [3].

Rishabh, Harmeet Singh Kathuria utilized multispectral spatial resolution imagery with spatial analysis Techniques solar generation on the rooftop has been calculated [4].

F. Jawaid and K. Nazir Junejo discussed the prediction of mean solar irradiation  $\text{Wh/m}^2$  various regression algorithms a comparative analysis of forecasting through artificial neural networks (ANN) against the standard regression algorithms [6].

Flavio Borfecchia, Emanuela Caiaffa, Maurizio Pollino used GIS and LiDAR Satellite Data to estimates of PV potential at roofs level [8].

Author R. Singh and R. Banerjee uses satellite imagery and Quantum GIS Software 1.8.0 and Pvsys tools to estimate macro-scale rooftop photovoltaic potential [7].

Teresa Santos, Nuno Gomes, Miguel Brito Evaluated the roof-top area suitable for installation of solar energy systems in the city of Lisbon, Portugal [9].

R. L. Bankert, M. Hadjimichael, A. P. Kuciauskas and et al. uses Satellite data from various sensors and platforms are being used to develop automated algorithms to assist in U.S. Navy operational weather assessment and forecasting [10].

### 3 Conclusions

A day by day the rules of environment pollutions are getting tougher, so it is need of solar plant installation. In this scenario various published research papers on the Geospatial and Machine Learning based method studied will be useful for applications for planning and development of solar power systems. Sun based radiation mapping, location assessment, and potential evaluation of solar energy for photovoltaic systems, specifically Machine Learning have been utilized for forecasting, such as regression techniques, Artificial Neural Network, Convolutional Neural Network model is suitable for accuracy at regional level.

Many researchers have used DSMs, DEM and radiation data is found useful and with the help of geospatial technology solar potential visualization, rooftop extraction, shading analysis: the effects of shadows on nearby light barriers such as trees and plants can be assessed, and spatial analysis tools for resolves such as simple geospatial analysis such as map algebra, overlaying, reclassification, slope, orientation, and distance. A radiation modeling allows forecast of the average daily and hourly global horizontal irradiation, beam radiation, and diffuse radiation to compute the insolation across whole scene or amount of sun light energy for given unit area to estimate the energy potential for planning of photovoltaic system installation.

**Table 1** For solar energy potential assessment review of the existing utilized data type/resources, techniques and findings to explore study of machine learning and geospatial technology

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
1	Yosoон Choi, Jangwon Suh and Sung-Min Kim, 2019	Applied Sciences	GIS-based solar radiation mapping, site evaluation, and potential assessment: a review	Solar map, satellite image, DEM, DSM, solar map, existing solar radiation map and r.sun (DEM)	Geographic information system based methods for site evaluation i.e. Boolean overlay, weighted sum, analytic hierarchy process, fuzzy and artificial neural network (ANN), shadow analysis, model validation are used	Radiation mapping, potential assessment and optimal site evaluation for solar power plants
2	Kakoli Saha, 2017	ICEGOV '17 SCII, March 07-09, 2017, New Delhi, AA, India © 2017 ACM. ISBN 978-1-45034930-7/17/03	A remote sensing approach to smart city development in India: case of Bhopal City, Madhya Pradesh	IRS-5 CARTOSAT-1 stereo images of Bhopal, acquired on 16th Feb 2012	Image preprocessing, automated extraction of urban roof area, production of reference map, accuracy assessment of the automated method, estimating potential for smart solutions	Rooftop solar PV, household level water harvesting, and green roofs can be ‘smart solutions’ addressing the core infrastructural problems of Indian cities

(continued)

**Table 1** (continued)

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
3	S. Sanders, C. Barrick, F. Maier and K. Rasheed, 2017	16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, 2017, pp 499–504	Solar radiation prediction improvement using weather forecasts	From Automated Environmental Monitoring Network (GAFMN). And National Oceanic and Atmospheric Administration's (NOAA) in the GRIB file format, its a compact binary format	Machine learning techniques in Weka	The improvement in accuracy while comparing machine learning methods in both one-hour and 24 h predictions shown here and indicates that the Random Forest ML model yielded with improvement
4	Xiaoyan Shao, Siyuan Lu, and Hendrik F. Hamann, 2016	2016 23rd International Workshop on Active-Matrix Flatpanel Displays and Devices (AMFPD), Kyoto	Solar radiation forecast with machine learning	Satellite images data	Machine learning based multi-model blending approach. Combining multiple meteorological models including numerous atmospheric and cloud prediction models based on satellite imagery as well as numerical weather prediction (NWP) products	Forecast based on solar irradiance

(continued)

**Table 1** (continued)

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
5	Rishabh, Harmeet Singh Kathuria, 2016	International Journal of Geomatics and Geosciences Volume 7, No. 2, 2016	Application of remote sensing and GIS technique in rooftop mapping and PV module layout design	(1) 1.28 m multispectral spatial resolution imagery (2) Multitemporal resolution digital imageries	RS and GIS spatial analysis techniques	Using photovoltaic layout, solar generation on the rooftop has also been calculated. This study proposed a method for significant reduction of time and cost involved in field visits for the installation of photovoltaic modules on a roof in any Rooftop Solar Power Project

(continued)

**Table 1** (continued)

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
6	F. Jawaaid and K. NazirJunejo, 2016	Sixth International Conference on Innovative Computing Technology (INTECH), Dublin, 2016, pp 355–360	Predicting daily mean solar power using machine learning regression techniques	National Solar Radiation Database (NSRD) National Climate Data Center (NCDC) National Oceanic and Atmospheric Administration (NOAA)	To predict the mean solar irradiation Wh/m <sup>2</sup> various regression algorithms (machine learning regression techniques) have been used in conjunction with various parameters related to solar irradiance	A comparative analysis of forecasting through artificial neural networks (ANN) against the standard regression algorithms. Therefore, demonstrate the ability of machine learning regression techniques, especially artificial neural networks to accurately predict the daily mean solar power up to a very high degree of accuracy

(continued)

**Table 1** (continued)

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
7	Flavio Borfecchia, Emanuela Caiaffa, Maurizio Pollino, 2014	European Journal of Remote Sensing—2014, 47: 195–216 Received 09/12/2013, accepted 27/01/2014	Remote sensing and GIS in planning photovoltaic potential of urban areas	LiDAR satellite data	RS and GIS techniques	Remote sensing techniques and advanced GIS modelling tools in order to support more realistic estimates of PV potential at roofs level in urban areas
8	R. Singh and R. Banerjee, 2013	IEEE 39th Photovoltaic Specialists Conference (PVSC), Tampa, FL, 2013, pp 2343–2347	Estimation of rooftop photovoltaic potential using satellite imagery and GIS	Google EarthTM-IKONOS and GE-1 Specifically from GeoEye's GE-1 satellite and weather data	Using satellite imagery and quantum GIS Software 1.8.0 and Pvyst tools	Estimating the macroscale rooftop photovoltaic potential
9	Teresa Santos, Nuno Gomes, Miguel Brito, 2011	Lena Halounová, Editor EARSeL, 2011	Solar Potential Analysis in Lisbon Using LiDAR Data	LiDAR satellite data	RS and GIS techniques	Evaluated the roof-top area suitable for installation of solar energy systems in the city of Lisbon, Portugal

(continued)

**Table 1** (continued)

Sr. No.	References/year	Name of journal/conference	Paper title	Data type/source	Techniques	Findings
10	R. L. Bankert, M. Hadjinichael, A. P. Kuciauskas, K. L. Richardson, J. Turk and J. D. Hawkins, 2002	IEEE International Geoscience and Remote Sensing Symposium, Toronto, Ontario, Canada, 2002, pp 708–710, vol. 2	Automating the estimation of various meteorological parameters using satellite data and machine learning techniques	Geostationary Orbiting Environmental Satellite (GOES) imagery. Special Sensor Microwave/Imager (SSM/I) data-satellite data	Supervised machine learning techniques are used to discover patterns in the data and develop associated classification and parameter estimation algorithms. These methods are applied to cloud classification in GOES imagery, tropical cyclone intensity estimation using <i>SSM/I</i> data, and cloud ceiling height estimation	Satellite data from various sensors and platforms are being used to develop automated algorithms to assist in U.S. Navy operational weather assessment and forecasting

## References

1. Y. Choi, J. Suh, S.-M. Kim, GIS-based solar radiation mapping, site evaluation, and potential assessment: a review. *Appl. Sci.* (2019)
2. K. Saha, A remote sensing approach to smart city development in India: case of Bhopal City, Madhya Pradesh, ICEGOV '17 SCII, 07–09 March 2017, New Delhi, AA, India, ACM, 2017. ISBN 978-1-4503-4930-7/17/03...\$15.00. <https://doi.org/10.1145/3055219.3055232>
3. X. Shao, S. Lu, H.F. Hamann, Solar radiation forecast with machine learning, in *23rd International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD)*, Kyoto, 2016, pp. 19–22. <https://doi.org/10.1109/AM-FPD.2016.7543604>. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7543604&isnumber=7543552>
4. Rishabh, H.S. Kathuria et al., Application of remote sensing and GIS technique in rooftop mapping and PV module layout design. *Int. J. Geomatics Geosci.* **7**(2) (2016). ISSN 0976-4380
5. S. Sanders, C. Barrick, F. Maier, K. Rasheed, Solar radiation prediction improvement using weather forecasts, in *16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Cancun, 2017, pp. 499–504. <https://doi.org/10.1109/ICMLA.2017.0-112>. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8260680&isnumber=8260597>
6. F. Jawaid, K. NazirJunejo, Predicting daily mean solar power using machine learning regression techniques, in *Sixth International Conference on Innovative Computing Technology (INTECH)*, Dublin, 2016, pp. 355–360. <https://doi.org/10.1109/INTECH.2016.7845051>. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7845051&isnumber=7845006>
7. R. Singh, R. Banerjee, Estimation of roof-top photovoltaic potential using satellite imagery and GIS, in *IEEE 39th Photovoltaic Specialists Conference (PVSC)*, Tampa, FL, 2013. <https://doi.org/10.1109/PVSC.2013.6744945>. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6744945&isnumber=6744078>
8. F. Borfecchia, E. Caiaffa et al., Remote sensing and GIS in planning photovoltaic potential of urban areas. *Eur. J. Remote Sens.* **47**, 195–216 (2014). [https://doi.org/10.5721/EuJRS2014\\_4713](https://doi.org/10.5721/EuJRS2014_4713)
9. T. Santos, N. Gomes et al., *Solar Potential Analysis in Lisbon Using LiDAR Data* (Lena Halounová, Editor EARSeL, 2011)
10. R.L. Bankert, M. Hadjimichael, A.P. Kuciauskas, K.L. Richardson, J. Turk, J.D. Hawkins, Automating the estimation of various meteorological parameters using satellite data and machine learning techniques, in *IEEE International Geoscience and Remote Sensing Symposium*, Toronto, Ontario, Canada, 2002, vol. 2, pp. 708–710. <https://doi.org/10.1109/IGARSS.2002.1025641>. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1025641&isnumber=22037>

# Framework for Faster Index Search



Chattar Singh and Vijay Gupta

**Abstract** *In the age of the web the quantity of archives is developing every day, every hour and every moment. In such a climate the recovery of the connected data is of main issue. In this paper the focal point of the idea of the full text search engine with the reference of Lucene and furthermore examine in respect of the indexing and framework plan.* In the framework, we have used the concept of the wordnet for the formation of the lexical chains, in order to find the related words and also to include the image search, we will be using the Tesseract API for extracting the text from the images, and then perform the search on that. This paper is only the conceptual explanation of the proposed work and the final implementation and result analysis we will include in the next research work.

**Keywords** Full-text search · Lucene · Indexing · Tesseract API

## 1 Introduction

Indeed, even with the immense measure of Internet data uncovered wherever and at whatever time, data recovery considering counterfeit is unmistakably not suits for the prerequisites of the current people. The search engine, based on the improvement of the full text recovery system and implying the rule of the organization programming structure, uses a specific PC program to assemble the data from the Internet. By then, it eliminates, examinations and methodology the data assembled. Thusly, an extraordinary UI can be formed and shows the recovery results to customers in order [1].

Standard unified search engine passes on all the limit modules of the system on the focal server. As needs be, the focal server load is excessively enormous, which

---

C. Singh (✉) · V. Gupta

International School of Informatics and Management, Jaipur, Rajasthan 302020, India  
e-mail: [chattarsinghr@gmail.com](mailto:chattarsinghr@gmail.com)

C. Singh

Rajasthan Technical University, Kota, Rajasthan 324010, India

will really impact. Thusly, the concentrated search engine never again deals with the social classes' solicitations in the hour of gigantic data. Also, the recovery eventual outcome of general search engine is wide and significant, consequently of which, customers scarcely find the data they need in confined time. Bit by bit directions to get the most exact search achieves the most restricted time has transformed into a main point of contention in the investigation of search engine.

The focal point of data search is considered as the full-text recovery innovation. The Full-text search innovation gave us the data recovery instrument as per the substance of data than the external sort of components one which are considering the combination of PC data such text, sound, picture as getting ready article [2]. Make every one of the possible terms in the rundown which are searched by network customers and furthermore cause individuals to manage and organize wide data and skilled organization customers to quickly and successfully recuperate any data they require.

## 2 Literature Survey

Many of the researches are done in the field of the full-text search. There have numerous execution utilizing different methodologies.

Cheng Lin et-al. 2016 Author breaks down the fundamental standard of the search engine and furthermore the strategy for Lucene to recreate indexing and searching, they builds an upward search engine dependent on Hadoop (HVSE) which tackles the issues identified with the customary brought together search engine and furthermore worked on low accuracy of the overall search engine. The work productivity of HVSE is substantially more than that of the incorporated search engine when managing huge measure of information. The design of HVSE comprises of the execution of the web based content mining, the web based structure mining, web utilization mining one which is through the strategies, for example, affiliation rules, arrangement, bunching, forecast model, succession design, way examination and so on [1].

Ruija Gao et-al. 2012 Lucene is the full text indexing engine toolbox which is written in Java, it support multi-client access, decreased admittance time and can be utilized cross-stage. This paper exhaustively investigates in profound the sort of working and also the execution of the Lucene, also the indexing and last the searching which are the involving the three-principle sort of modules from framework engineering and contrasted the Lucene full text search and the String recovery's reaction time, and the trial results then shows that the concept of Lucene full text search has also the quicker sort of recovery speed [2].

Khaled Nagi et-al. 2016 In this paper, creator analyzes the huge scope search engine which can be applied on cloud. Creator utilizes the open-source libraries, including Lucene, Solr, mongoDB, Cassandra, and Hadoop. They explore by means of information parceling, excess through replication, consistency either through the NoSQL data sets or through open-source synchronization libraries, like Zookeeper. They develop a little and a huge group. In their investigations, they measure both the

speed of indexing just as the search time and the throughput of the searching strings. What's more, the consequences of the examinations show that Solr and Hadoop give the best tradeoff as far as versatility, dependability and sensibility [3].

Sarah Aldawood et-al. 2016 To test the search execution creators has fabricated a basic program utilizing the Lucene library and Java language. The program is fit for deciphering distinctive document designs and will keep up with the index of the records gave as an info. The search cycle starts with the contribution of word from the client and as the outcome the program will show the name of the record which contains the necessary work. The program utilizes the Lucene library, first and foremost to play out the indexing system for the gathering of documents. What's more, furthermore, the consequences of searching text that the searching system recovery on various inquiries. The basic trial by the creator utilizing the basic program showed that Lucene fulfilled speed execution in the case of both the indexing and also the searching cycles [4].

Mohit Bhansali et-al. 2013 In this paper creator present a search framework which is made out of four modules for example indexing system, dissecting text, index putting away and searching instrument. Aside from these modules creator give execution of UI, and related modules associate with one another. This paper presents the construction of work area search engine framework and its parts incorporate indexing, dissecting, index putting away and searching. The creator has tried the indexing and searching capacities of the system [5].

X. Chen and L. Xu 2016 in this paper an ERM framework is proposed which depends on the Lucene and on the subject index. It is the data recovery framework intended for instruction asset recovery. Here the subject model is utilized which contains the points index, to play out the full-text search. Creators made a new dataset for the testing of the framework utilizing Baidu Wenku dataset [6].

F. Ramli et al. 2016 present a philosophy-based methodology to plan and furthermore for fostering the new portrayal for IR framework to defeat the hardships of the traditional catchphrase based ap-proach. The records for framing the dataset are the verifiable reports identified with the Vietnam War. Authors utilized this methodology of Simple News and Press Ontologies (SNaP) cosmology and expanded the methodology using the custom jargon identified with the verifiable area. With the end goal of result investigation, the contrasted the presentation and BM25 IR model dependent on corpus of 133 reports from Wikipedia [7].

A. I. Ismail et al. 2017 presents the informatics portrayal in regards to the search engine, how the information is principle trained and streams in it. Comparable, likewise clarify the most common way of obtaining the information and afterward to recover the information utilizing the Search Engine [8].

S. Lakhara and N. Mishra 2017 creators proposed the two searching calculations one for the catchphrase searching and one more for the record searching, which are enlivened by Lucene search and it is utilized to perform search on pdf, doc and so forth reports and the area of the search is the neighborhood hard plate [9].

I. Safder and S. Hassan 2018 creators proposed the profound learning-based framework one which improves the capability of the search instruments by the idea of the grouping the calculation explicit metadata, like the exactness, the accuracy and

furthermore the review, and the methodology utilizes the regular terms as in ‘sack of words’ models [10]

J. Jabbar et al. 2019 the proposed DMS one which contains the electronic administrator board, the customer application, and furthermore the portable application. The outcomes additionally have demonstrated that the proposed idea of searching system is a lot quicker and furthermore has likewise has further developed text recovery one that will help clients to search effectively and furthermore the recovered data quickly [11]. J. YU et al. 2019 the proposed work dependent on the idea of topical learning as the means and more append more significance is given to such three of the weight markers as the news title, the length and furthermore idealness to update the weight work based on the Lucene full-text recovery calculation [12].

Z. Youzhuo et al. 2020 proposed a full-text recovery framework dependent on the engineering of Lucene, which is then incorporated with the Hadoop design to build up the dissemination network large server farm, one that could give the artistic and furthermore the full-text inquiry across all dispersion network oversee frameworks [13].

### 3 Full-Text Search

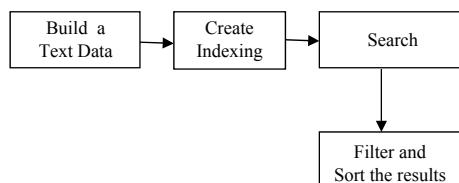
The concept of the Full Text Search is divided into following four segments or steps.

The Fig. 1. Explains the steps which are performed using the full-text search.

#### 3.1 Build a Text Database

In the process of the Full-Text search, the first step is to create a Text Database on which store all the required information which can be retrieved by the user. The Text Database should be identifiable and must have the low level of data redundancy. The model of the Text database is created once and it should not change once we have finalized it [14].

**Fig. 1** Full-Text Search Concept



### ***3.2 Create Indexing***

The second step is to create the index with the models all depending upon the text in the database. Indexing is required for efficient and the quick retrieval of the information from database. And the way the indexing is done or the way we create the index is all rely on the scale or extent of the Information based retrieval systems. If we consider the large-scale information-based retrieval systems like the Google, the Baidu, they make the use of the concept of the inverted index [15].

### ***3.3 Search***

The third step, after we have created the text database and index, is to perform the search. The process of search begins by capturing the data input by user and preprocessing the data which is being entered and on the corresponding to it will return back the required information using the concept of the information retrieval system [15].

### ***3.4 Filter and Sort the Results***

Once the search process is over the information retrieval system will return the result to user after filtering and sorting it according to a specific so the user will get the relevant information and in the proper manner [15].

Lucene is created by the Apache Software Foundation and it is open-source full text search engine toolkit. It is not a full-text based search engine but even though it is a framework which enables the users to make use of query and text indexing engine. It is quite simple and by making the use of this API interfaces; search engine can be quickly developed.

## **4 Full-Text Search Engine of Lucene**

### ***4.1 Module Structure of Lucene***

- Lucene framework composed of the seven modules out of one which the five of the important modules are as follows :
- Org.apache.lucene.analysis This module is devoted to perform the segmentation of the document and to remove the stop words which play no role in the information search. This module further composed of two sub-modules, namely StandardAnalyzer and SimpleAnalyzer.

- Org.apache.lucene.document This module is mainly then responsible for the proper management of the fields and these fields are further divided into the text and the date fields.
- Org.apache.lucene.index This module is responsible for the management of index, insertion and deletion of records. The process goes by marking the terms associated with the index so the only the required portion of the data is directly accessed rather than going through the whole data.
- Org.apache.lucene.search This module is responsible for the retrieval of the information on the basis of the search query.
- Org.apache.lucene.queryParser queryParser, This module with parse the query and according hand over the search module.

## ***4.2 The Indexing of Lucene***

In the indexing of Lucene, it provides the classes IndexWriter and DocumentWriter. IndexWriter is concerned with the concept of adding documents to the indexing and DocumentWriter deals with the writing process of indexing.

## **5 Proposed Concept Overview**

### ***5.1 Methodology Proposed***

There are many different concepts and modules which are involved in the process and are explained as follows:

- Lucene is a full-text search library for Information Retrieval (IR). Content can be added to Lucene from different sources, similar to a SQL information base, a filesystem, or even from sites. It gives search over reports; where an archive is basically an assortment of fields. A field comprises of a field name that is a string and at least one field type. Fields are obliged to store just a single sort of information, either parallel, numeric, or text information. There are two methods for putting away text information: string fields store the whole thing as one string; text fields store the information as a progression of tokens. Lucene has an exceptionally expressive search API that takes a search inquiry and returns a bunch of reports positioned by pertinence with archives generally like the question having the most elevated score.
- WordNet is a lexical information base of semantic relations between words in excess of 200 dialects. WordNet joins words into semantic relations including equivalents, hyponyms, hypernym, homonym and meronyms. Words from the very lexical classification that are generally equivalent are assembled into synsets. All synsets are associated with different synsets through semantic relations.

- A lexical chain is a succession of related words like equivalents, hyponyms, hypernym, holonym and meronyms recorded as a hard copy, spreading over short (contiguous words or sentences).
- Tesseract is an open-source text acknowledgment (OCR) Engine, accessible under the Apache 2.0 permit. It very well may be utilized straightforwardly, or (for developers) utilizing an API to separate printed text from pictures. It upholds a wide assortment of dialects. The Tesseract OCR exactness is genuinely high out of the crate and can be expanded altogether with a very much planned Tesseract picture preprocessing pipeline

## 5.2 *Procedure Adopted for Proposed Work*

- In the proposed system, the first input is the expression, then the expression is split into the words and synonyms, hyponyms, hypernym, homonyms and meronyms are found for the word. For the process of the related words and lexical chain finding will be done using the wordnet.
- Then the search operations involve first stage processing using
  - a. Query Parser
  - b. Analyser
- Now the documents which are taken for analysis includes, text documents, word documents, excel files, pdf, images and HTML (Documents extensions: doc, docx, html, pdf, ppt, pptx, xml, xls, xlsx, jpeg etc.)
- For the process the image containing the text we will use the tesseract OCR API for the conversion of the image into to text and then apply the process of
  - c. Index Searcher
  - d. Index Reader
- After that it will return the directory location of top docs as search result having filename, keyword and line number in which that keyword appears.

## 6 Conclusion

All in all, because of number of various sort and the arrangements accessible for the electronic records it is not difficult to search the text in the report. Number of searches search engine based applications are accessible which are very valuable during the time spent the full-text search, yet assuming we look at the Lucene search engine library is by all accounts one of the most mind-blowing choices to play out the speedy search and recovery of the required and significant data. Here, we have only introduced what will be our proposed work or what model we will be going

to build, in the coming research paper we will simulate the proposed approach and compare it with the current research works.

## References

1. C. Lin, M. Yajie, Design and Implementation of vertical search engine based on Hadoop, Eighth international conference on measuring technology and mechatronics automation (2016)
2. R. Gao, D. Li, W. Li, Y. Dong, Application of full text search engine based on Lucene, in *Advances in Internet of Things*, SciRP (2012)
3. K. Nagi, in *Bringing Search Engines to the Cloud using Open Source Components* (IEEE, 2016)
4. S. Aldawood, in *Text Analyzer: Search Text Based on Lucen Search Engine Library* (IEEE, 2016)
5. M. Bhansali, Searching and analyzing qualitative data on personal computer, IOSR J. Comput Eng
6. X. Chen, L. Xu, An educational resource retrieval mechanism based on Lucene and topic index, 2016 13th web information systems and applications conference (WISA), Wuhan, pp. 125–128 (2016)
7. F. Ramli, S.A. Noah, T.B. Kurniawan, Ontology-based information retrieval for historical documents, 2016 third international conference on information retrieval and knowledge management (CAMP), Bandar Hilir, pp. 55–59 (2016)
8. A.I. Ismail, R. Kandan, E.M. Goortani, M.N.M. Mydin, M.F. Khalid, O.H. Hoe, Reference architecture for search infrastructure, 2017 7th IEEE international conference on control system, computing and engineering (ICCSCE), Penang, pp. 115–120 (2017)
9. S. Lakhara, N. Mishra, Design and implementation of desktop full-text searching system, 2017 international conference on intelligent sustainable systems (ICISS), Palladam, pp. 480–485 (2017)
10. I. Safder, S. Hassan, DS4A: deep search system for algorithms from full-text scholarly big data, 2018 IEEE international conference on data mining workshops (ICDMW), pp. 1308–1315 (2018)
11. J. Jabbar, W. JunSheng, L. Weigang, I. Urooj, Implementation of search engine with Lucene in the document management system, 2019 IEEE 2nd international conference on electronics and communication engineering (ICECE), pp. 1–4 (2019)
12. J. Yu, A. Su, W. Liu, X. Cheng, J. Yang, Thematic learning-based full-text retrieval research on BRITISH and American journalistic reading, 2019 14th international conference on computer science & education (ICCSE), pp. 611–615 (2019)
13. Z. Youzhuo, F. Yu, Z. Ruifeng, H. Shuqing, W. Yi, Research on Lucene based fulltext query search service for smart distribution system, 2020 3rd international conference on artificial intelligence and big data (ICAIBD), pp. 338–341 (2020)
14. A. Arslan, O. Yilmazel, Quality benchmarking relational databases and Lucene in the TREC4 adhoc task environment, Proceedings of the international multiconference on computer science and information technology, Wisla, pp. 365–372 (2010)
15. L. Qian, L. Wang, An evaluation of Lucene for keywords search in large-scale short text storage, 2010 international conference on computer design and applications, Qinhuangdao, pp. V2–206-V2–209 (2010)

# Image Classification by Optimized Convolution Neural Networks



Eva Tuba , Ira Tuba , Romana Capor Hrosik , Adis Alihodzic , and Milan Tuba

**Abstract** Considering the fact that digital images are used in almost all scientific areas and they are a big part of everyday life, it is obvious that the importance of good methods for processing and analyzing them is great. One of the most frequent tasks in various applications that use digital images is image classification. A revolutionized improvement in this area was achieved with convolutional neural networks (CNN). The convolutional neural networks managed to achieve classification accuracy significantly better compared to previously proposed and used methods. Even better results can be obtained by tuning CNN hyperparameters. Since this is a hard optimization problem, swarm intelligence algorithms can be successfully used. In this paper, we propose bare bones fireworks algorithm for tuning a selected subset of hyperparameters and it was tested on the benchmark dataset for handwritten digit recognition, MNIST. The proposed method achieved higher classification accuracy compared to the methods from the literature.

**Keywords** Image classification · Convolutional neural networks · Optimization · Metaheuristic · Swarm intelligence · Hyperparameter tuning

---

E. Tuba · I. Tuba · M. Tuba (✉)  
Singidunum University, Belgrade, Serbia  
e-mail: [tuba@ieee.org](mailto:tuba@ieee.org)

E. Tuba  
e-mail: [etuba@ieee.org](mailto:etuba@ieee.org)

I. Tuba  
e-mail: [ituba@ieee.org](mailto:ituba@ieee.org)

R. C. Hrosik  
University of Dubrovnik, Dubrovnik, Croatia  
e-mail: [romana.capor@unidu.hr](mailto:romana.capor@unidu.hr)

A. Alihodzic  
Faculty of Science, University of Sarajevo, Sarajevo, Bosnia and Herzegovina  
e-mail: [adis.alihodzic@pmf.unsa.ba](mailto:adis.alihodzic@pmf.unsa.ba)

## 1 Introduction

Digital images are included in applications used in various areas such as medicine [5], autonomous vehicles [4], security [11], and others. One of the main tasks in those applications is digital image classification. Considering the wide use and importance of image classification in these applications the improvement of the obtained classification accuracy is in high demand. Brings great benefits in all these areas. For decades, researchers have been working on finding the best possible image processing algorithms that extract different features that are used as inputs for various classifiers. These features are chosen to best describe the regions of interest in order to make the classification as accurate as possible. Some of the examples of often used features are edges, texture, invariant moments, contours, etc. Nowadays all of that is not necessary due to convolutional neural networks (CNN). Technology development and great computation power have made usage of CNN possible and made a revolution in digital image classification. The advantage of the CNN compared to other classification methods is that no pre-processing of the image and feature extraction is needed. Without all that the CNN achieve significantly better accuracy.

One of the issues with the CNN is tuning its hyperparameters. Numerous hyperparameters from the large range of possible values makes this problem a hard optimization problem. In the past decades, swarm intelligence algorithms (SI) have been widely applied to this kind of problem and achieved good results. Applications of the SI algorithms to the problem of CNN hyperparameters tuning is a new research area and there is a lot of space for improving the current state-of-the-art results. In this paper, we proposed bare bones fireworks algorithm for tuning a subset of CNN hyperparameters.

The rest of the paper is organized as follows. Section 2 gives details about the convolutional neural networks. In Sect. 3, an application of the SI algorithms for CNN hyperparameter tuning along with the proposed method were presented. Simulation results and comparison with the methods from literature are given in Sect. 4. Conclusion and future work were presented in Sect. 5.

## 2 Convolutional Neural Networks

Classification of digital images is necessary for many different applications and lately, it is done by convolutional neural networks. In the last decade, CNN were applied to various problems such as medical image analysis [3], satellite image analysis [10], wind power forecasting [12], noise removal [7], etc. Implementation and training of a CNN is a relatively simple task due to the help of numerous available software solutions. The main advantage of the CNN is that it will, in most cases, by far outperform other machine learning methods. Another thing that sets CNN apart from other methods, in addition to great results, is that no prior image preparation

is required, raw data can be used as the input. Usually resizing the image is the only preparation done before applying the CNN.

What differentiates the CNN from simple artificial neural networks is the specific way of input processing. Instead of treating inputs separately, as noncorrelated values, the CNN perform the convolution operation on neighboring inputs. This allows the network to make some conclusions based on the correlation in input data which is very important for processing digital images or any other signal that has spatial correlation like voice signal or EEG. This type of analysis is done by a special type of layer named the convolution layer. Besides convolutional layers, there is a pooling layer that is used to reduce the complexity of the network. In the pooling layer, the matrix dimension that represents the output of the previous layer is reduced by taking the average of the pixels in certain sub-matrices (i.e. average pooling layer) or taking the maximal value as a representative of the sub-matrix (max-pooling layer). After the input layer, one or several convolutional layers can be followed by the pooling layer and at the end of the network, before the output layer, there are one or more fully connected or dense layers. This is a typical CNN architecture.

Each of the mentioned layers has specific hyperparameters that should be tuned. Each convolutional layer has kernel size, stride, padding, number of output feature maps as hyperparameters. The pooling layer has to have kernel size, padding, stride and type defined. Standard hyperparameters of the ANN are also part of the CNN: learning rate, batch size, number of epochs, optimization algorithm, activation function, etc. Besides that, some hyperparameters consider the CNN architecture. For example, the number and order of convolutional and pooling layers, number of dense layers and number of nodes in each one of them. This leads to the conclusion that despite all advantages of the CNN, there are challenges and hyperparameters tuning is one of the main ones. The numerous hyperparameters that should be tuned for each problem along with the large ranges of the possible values for each of them, make this problem a type of combinatorial problem that cannot be solved by any deterministic method in a reasonable time.

Tuning hyperparameters of the classifier is a problem and for solving it different methods were proposed. The simplest but not very efficient methods are random or grid search as well as estimating good values based on the previous experience and knowledge, i.e. guestimating. For these kinds of problems, different metaheuristics have been successfully applied in the past. One of the classes of optimization methods that showed good results for tackling similar problems is swarm intelligence algorithms (SI). Application of the SI algorithms for the CNN hyperparameters tuning have been studied in recent years but further research is needed [1, 2].

### 3 Swarm Intelligence for CNN Optimization

Finding the optimal CNN for the image classification is an important task and it is a hard optimization problem. Using swarm intelligence algorithms for tackling this problem is a promising method considering the success of the SI for other similar

problems. Optimization by the SI algorithms is done by using a swarm of simple agents that has operators for moving them, sending and receiving information about their positions. The main parts of any SI algorithm are exploration and exploitation operators that define the movement of each member of the population. The exploration is used for searching the whole space of possible solutions with an aim to find areas where the optimal solution is. After finding the good areas of the search space, the exploitation operator is used for fine search. The important part is to have a balance between these two operators, which means to ensure fast and accurate convergence toward the optimal solution.

Differences between SI algorithms are in the used operators and usually, some parameters control the balance between exploration and exploitation that should be set for each problem individually, either empirically or by using some heuristic. The complexity of the algorithm is highly dependent on these two operators.

The CNN hyperparameter tuning is rather a difficult optimization problem, not just because of the number of hyperparameters but also because of the fitness function. The most logical fitness function is classification accuracy. In order to obtain classification accuracy, it is necessary to train the whole network. Since the CNN are usually deep networks, the training is rather time consuming and performing several thousand fitness function evaluations would lead to losing the advantage of using the SI algorithm. This can be resolved by transforming the solution representation or by adapting different fitness functions. In this paper, the focus was on obtaining a good CNN by a small number of fitness function evaluations. The only adaptation of the SI algorithm was for integer solution values. Considering the nature of the selected hyperparameters, the only adaptation was rounding obtained solutions to the nearest integer.

Due to the complexity of the problem, we choose the simple SI algorithm, bare bones fireworks algorithm.

### ***3.1 The Proposed Method for Tuning CNN Hyperparameters***

Bare bones fireworks algorithm (BBFWA) was proposed in 2018 by Li and Tan [9]. The BBFWA is a minimalist swarm intelligence optimization algorithm. The exploration and exploitation are controlled by changing the size and position of the search space. New solutions are always generated around the previously best solution. The space around the best solution is increased by a factor  $C_r$  if exploration is needed which is when a better solution is found in the previous generation. It is considered that if a better solution is found, then a wider search is needed. On the other hand, if the best solution remains the same, exploitation is needed around that solution so the search space around it where the new solution will be generated in the next generation is reduced for factor  $C_a$ . Initially, the hypercube where the solutions are generated should cover the whole search space, so the size of the hypercube is equal to the  $A = Ub - Lb$  where  $Ub$  and  $Lb$  are upper and lower bound of the search space. The center of the hypercube is the previous best solution.

Another reason why the BBFWA is a good choice for the CNN hyperparameters tuning is the fact that most SI algorithms are designed for real solutions and exploration and exploitation operators are designed to slightly move already generated solutions while the BBFWA depends only on the previous best solution. This is important since in the case of hyperparameter tuning, a small change in for example number of convolutional layers leads to a completely different network that can give significantly different classification accuracy and the benefits of carefully designed operators lose the purpose. In the BBFWA, new solutions around the previous best solutions are generated, so only promising areas are kept through generations, but the search of this area is not dependent on the quality of the solutions from the previous generation. For the BBFWA, adaptation for the integer solutions is simple rounding obtained real values.

In this paper, we used the BBFWA to find the optimal values for some CNN hyperparameters. The subset of hyperparameters searched by the BBFWA is listed in Table 1.

The solution representation takes into account that the number of convolutional layers cannot be larger than 3, so there is one variable determining how many convolution layers are used. Their values are 0 or 1, i.e. the convolutional layer is used or not. For each of these 3 layers, there are 4 variables, for each convolutional hyperparameter. Fully connected (FC) layers are limited to 2. Similar to the convolutional layers, for each potential FC layer, 4 variables are made. In the end, 3 variables are used for batch size, learning rule and learning rate. This means that the total number of variables is 25. This is a high dimensional problem and usually, for such problems, a large number of fitness function evaluations is needed. However, the fitness function in this paper is classification error and for finding that, training a complete

**Table 1** Convolutional neural network hyperparameters

Category	Hyperparameter	Notation
Convolutional layer	Number of convolutional layers	$n_c$
	Number of kernels	$n_f$
	Kernel size	$f_s$
	Activation function	$a_c$
	Pooling layer size	$p_s$
Fully-connected layer	Number of fully-connected layers	$fc_s$
	Connectivity pattern	$cp$
	Number of units	$n_u$
	Weight regularization	$wr$
	Dropout	$d$
General hyperparameters	Batch size	$b_s$
	Learning rule	$lr$
	Learning rate	$\alpha$

CNN is needed. This is rather an expensive operation and it is not possible to have several thousand fitness function evaluations. In order to reduce the time needed for fitness function, we trained the CNN for only 10 epochs. Additionally, we empirically concluded that solutions that have classification errors larger than 10 after 5 epochs will not achieve the desired classification accuracy, so the training is stopped after only 5 epochs. In [13] limitation to 10 epochs was also used and additionally, the networks were trained with only 50% of data so we also added this limitation in the fitness function. The 50% data used for training is randomly chosen. After finding the best values for hyperparameters, the CNN is trained with all training data through 40 epochs and tested on a test set.

## 4 Simulation Results

Implementation of the proposed method was done in Python 3.8 with the usage of the PyTorch package for the CNN. Platform used for simulations was with Intel® Core™ i7-11700 K CPU at 5 GHz, 16 GB RAM and NVIDIA RTX 2060 graphic card. The OS is Windows 11 Professional OS.

The proposed method was compared with the method proposed in [13]. In [13], the CNN hyperparameters were optimized by the artificial bee colony (ABC). Considered hyperparameters were number of convolutional layers and number of feature maps for each one, kernel size, activation function, pooling size after each convolutional layer, number of fully connected (FC) layers, number of nodes in each FC layer, weight regularization, dropout probability, batch size, learning rule and learning rate. They introduced a specific coding of the solution and reduced the fitness function evaluation time by training the network only for 5 epochs and using only 50% of data for training in each epoch. The number of fitness function evaluations was 4,000. The proposed method was tested on the MNIST dataset [8], a standard benchmark dataset with images of handwritten digits. Images are a gray scale of size  $28 \times 28$ . Dataset has 60,000 images for training and 10,000 images in the test set.

Parameters of the BBFWA were set as follows:  $C_a = 0.6$  and  $C_r = 1.4$ , population size was 20 and the maximal iteration number was 40, which means 800 fitness function evaluations. Considering the fact that our proposed fitness function takes 10 epochs for training the CNN and in [13] 5 epochs, the number of our fitness function evaluations is equivalent to 1600 evaluations in [6].

Obtained results by the proposed method along with the results reported in [13] are presented in Table 2.

The proposed method found hyperparameters of the CNN that give a lower classification error compared to the ABC method proposed in [13] and methods that were used for comparison in [13]. The proposed BBFWA achieved the average classification accuracy on the test set of 99.34% while the second best method was the ABC method proposed in [13] where the average classification accuracy was 99.26%. The number of fitness function evaluations was 2.5 times smaller in our proposed method compared to the ABC method [13] but due to the nature of the BBFWA algorithm, it

**Table 2** Classification error of the proposed method and results from [13]

	Algorithm	best	worse	mean
EXACT	1.57	4.12	2.80	
Alejandro	0.48	2.78	1.43	
LeNet-5	0.82	1.20	1.07	
ABC	0.62	0.98	0.74	
BBFWA	<b>0.58</b>	<b>0.82</b>	<b>0.66</b>	

Bold significance are better results

managed to find better solutions. Limiting the size of the promising areas and guiding toward them turns out to be better for exploration and exploitation operators for the problem of tuning CNN hyperparameters than limiting new solutions by the previous ones and allowing only small movements. Based on the results, it can be concluded that more freedom in generating new solutions leads to better classification accuracy.

## 5 Conclusion

This research proposes the usage of the bare bones fireworks algorithm for the CNN hyperparameter tuning. The proposed method is used for finding optimal values for 8 hyperparameters. Obtained results were compared to the results from the literature. Considering the classification accuracy, the proposed method outperformed the ABC method. Moreover, the BBFWA was able to find better CNN with fewer fitness function evaluations, which is important for this problem since this is the most expensive operation. Future research will add more hyperparameters and consider different datasets. Modification of the solution representation and fitness function are also topics of interest for further work.

## References

1. N. Bacanin, T. Bezdan, E. Tuba, I. Strumberger, M. Tuba, Monarch butterfly optimization based convolutional neural network design. *Mathematics* **8**(6), 936 (2020)
2. N. Bacanin, T. Bezdan, E. Tuba, I. Strumberger, M. Tuba, Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics. *Algorithms* **13**(3), 67 (2020)
3. T. Bezdan, M. Zivkovic, E. Tuba, I. Strumberger, N. Bacanin, M. Tuba, Gliomabrain tumor grade classification from mri using convolutional neural networks designed by modified FA. In: *International Conference on Intelligent and Fuzzy Systems* (Springer, 2020), pp. 955–963
4. H. Gao, B. Cheng, J. Wang, K. Li, J. Zhao, D. Li, Object classification using cnn-based fusion of vision and lidar in autonomous vehicle environment. *IEEE Trans. Industr. Inf.* **14**(9), 4224–4231 (2018)

5. R.C. Hrosik, E. Tuba, E. Dolicanin, R. Jovanovic, M. Tuba, Brain image segmentation based on firefly algorithm combined with k-means clustering. *Stud. Inform. Control* **28**(2), 167–176 (2019)
6. D. Karaboga, B. Basturk, On the performance of artificial bee colony (abc) algorithm. *Appl. Soft Comput.* **8**(1), 687–697 (2008)
7. H.Y. Khaw, F.C. Soon, J.H. Chuah, C.O. Chow, High-density impulse noisedetection and removal using deep convolutional neural network with particle swarm optimisation. *IET Image Proc.* **13**(2), 365–374 (2019)
8. Y. LeCun, C. Cortes, MNIST handwritten digit database (2010). <http://yann.lecun.com/exdb/mnist/>
9. J. Li, Y. Tan, The bare bones fireworks algorithm: A minimalist global optimizer. *Appl. Soft Comput.* **62**, 454–462 (2018)
10. M. Segal-Rozenhaimer, A. Li, K. Das, V. Chirayath, Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (CNN). *Remote Sens. Environ.* **237**, 111446 (2020)
11. Y. Xu, Z. Zhang, G. Lu, J. Yang, Approximately symmetrical face images for image prepro- cessing in face recognition and sparse representation based classification. *Pattern Recogn.* **54**, 68–82 (2016)
12. X. Yang, Y. Zhang, Y. Yang, W. Lv, Deterministic and probabilistic wind power forecasting based on bi-level convolutional neural network and particle swarm optimization. *Appl. Sci.* **9**(9), 1794 (2019)
13. W. Zhu, W. Yeh, J. Chen, D. Chen, A. Li, Y. Lin, Evolutionary convolutional neural networks using ABC. In: *11th International Conference on Machine Learning and Computing* (ACM, 2019), pp. 156–162

# Review: Benefits of Mining Fashion Data



Monika Sharma and Priyanka Verma

**Abstract** Data mining is said to be an action that makes it useful for extracting usable data from a big set of many types of unfinished data. It is used to analyze patterns of data in huge batches of data using the software. In the fashion industry, data mining is used by, fashion designers to see the very popular colors and modify or refashion their designs that satisfy the requirement of customers. Designers design clothing according to the requirements of customers. They use fashion data to set the price of clothing. The fashion industry can use data mining to view popularity trends on a granular level of all, seeing who will buy what and why? With information mined, labels can divert efforts towards focusing the products with nascent potential.

**Keywords** Fashion industry · Data mining techniques · Data mining strategies · Software used

## 1 Data Mining

Data mining is the mining of facts gathered from different sources. So many organizations collect a variety of data to make wonderful decisions and without proper facts and figures, it's not possible for any organization to make the best decisions and so data and facts are collected from time to time from different sources before a new product launch, It's a need of every organization to collect data due to product demand, some time when customer preferences, sometimes competitors and sometimes when the data is not collected beforehand when the organizations launch a product it can lead to failure any number of ways spread, like less of demand and not able to meet the need of the customer. Data mining is also said to be an action

---

M. Sharma (✉) · P. Verma  
IIS (Deemed to Be University), Jaipur, India  
e-mail: [monikamissrajasthan@gmail.com](mailto:monikamissrajasthan@gmail.com)

P. Verma  
e-mail: [priyanka.gianchandani@iisuniv.ac.in](mailto:priyanka.gianchandani@iisuniv.ac.in)

that makes it useful for extracting usable data from a big set of many types of unfinished data [1]. It is used to analyze patterns of data in huge batches of data using the software.

### ***1.1 Knowledge Discovery in Data (KDD) is also Called Data Mining***

#### **Process of Data Mining**

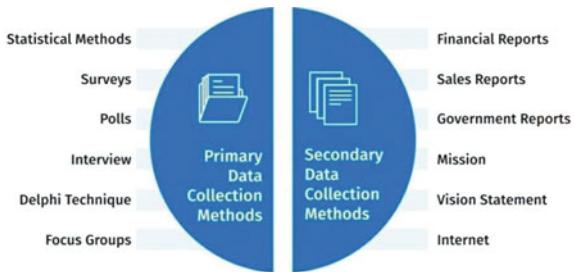
The process of data mining is based on the following phase:

- Cleaning of data—Very first step in the field of data mining. It removes all noise and data that is incomplete from the collection filling the missing data, removing the noisy data.
- Integration of data—When multiple data sources like data cubes, databases, or files are combined for analysis, this type of process is called the integration of data. It can help in improving the speed and accuracy of the data mining process
- Reduction of data—This is a technique that is applied to obtain data that is relevant for analysis from the collection of data. An example we take such as the no of attributes in the dataset reducing, and the original data volume by a smaller form of data representation which is replacing.
- Transformation of data—Transformed of data into a suitable form for the process of data mining. Some examples are—remove noise from data using regression and clustering.
- Mining of data- It is the process of identifying patterns that are interesting and knowledge from a large amount of data.
- Evaluation of Pattern-Summarization and visualization of data methods to make the data understandable by the user.
- Representation of knowledge—Some of the data visualization tools used to represent the fine mined data in the form of tables, reports, etc.

### ***1.2 Data Mining Methods***

- Primary data collection-This data is collected with a completely new experience like first-hand experience or defiantly never used in past.
- Secondary data collection-This data has been used somewhere in the past, it is very less time-consuming, and very easily available (Fig. 1).

**Fig. 1** Primary or secondary data collection methods of data mining [2]



## 2 Data Mining in the Fashion Industry

In the Fashion industry, data analysis is done through data mining. It is the future of analyzing the market and being aware of upcoming trends to hit the market with the right product at the right time. Being aware of upcoming trends using Data Analysis enables the brands to design the right product for the targeted customers. Data analysis of data mining enables the Fashion Industry brands to stay aware of the upcoming trends, cutting the operational costs and meeting the expectations of the consumers. You can get help from Data Mining from past trends to make sure that the design you are going to introduce will be a success and will face the market trends. It also plays a key role in predicting the future and for how long a trend is expected to stay in. This ensures that you design and products that do not stain your brand image. This also enables you to be prepared in advance to change the design and compete with the market to keep your products trending.

- Detecting and predicting fashion trends using data mining.
- Fashion brands and new start-ups are actively engaging with data mining.
- Data mining is very important for the fashion industry
- Data mining can give us the desire of customers.
- Data mining methods make work very easy.
- By using data mining techniques, we can increase the sales of your company very much.

We can know what is new fashion trends going on in the market what do people need? with the help of data mining.

### 2.1 Problems Faced in the Fashion Industry

Clothing and fashion production have numberless provocations to achieve. Some of them are:

- Digitalization-For shoppers when the industry launches some better conception to swap the way of shopping. But somewhere production breaks down. The customer

shopping choices or behaviors of the consumer are the deeper issues that the industry is facing to fill the gap between fashion and technology, brands, and retailers need a digital mindset. Fashion production is more focused on long-lasting technology and real transformation.

- Consumer shift-It is too hard for many market competitors to put upgrade with the present trends, market needs, or consumers. Every time the change adapted with time is most important.
- New generation consumer-As many of all companies are struggling to catch and keep the user's attention through the social networks and same platforms, it become very important for such fashion companies to rethink. Content should be developed on the platform where it will be published so far and the market to which it is directed that facilitates the purchase all over.
- New materials-Many companies survey better substitutes to traditional materials, deciding to replace them in their collections with all better possible and more fantastic fabrics.

## ***2.2 Use of Data Mining in the Fashion Industry***

Techniques that are machine learning, data mining, and artificial intelligence are used to identify garments categories and subcategories. Given below is how data mining can be applied in the fashion industry.

- Apparel industry's result data is mined, to get information related to clothing, fabric, print, trim, color, and shape.
- Classification techniques such as naïve bayes, decision trees, Bayesian Forest, and random forest were applied to deep fashion.
- FOREST classifier was comparatively better for the categorization of garments, upper body garments, lower body garments, and whole-body garments.
- ASSOCIATION ANALYSIS-discover association patterns or correlation patterns existing among the clustered group.
- SURVEY QUESTIONNAIRE was designed to demonstrate how we use a rule of association to mining automatically extract the amazing patterns from the large data collected by survey.
- RAPID MINING used for relevance clustering analysis.

## **3 Data Mining to Analyze Fashion Consumers**

This section discusses how various data mining techniques can be used to analyze fashion consumers.

The clustering data mining method helps to identify 3 categories of vogue adopters.

- 1 Fashion leaders-Front runners we can say early adopters, people who are very daring, bold, and fearlessly involved in up-to-date (newborn) fashion.

- 2 Fashion followers-Before the usual time people, those who follow recently developed fashion.
  - 3 Fashion laggards-Behind the time majority of people, those who never adopt or are at the back of the line to take on new trend fashion.
- Association analysis—Association analysis can be performed on Three clustered groups Fashion leaders, followers, laggards, to find association and correlation patterns existing between these them
  - Survey questionnaire—was designed in such a fact-finding study to make use of association rule mining to spontaneously take out all interesting patterns from the huge collection of survey data.
  - Rapid mining—used for relevance clustering analysis.

## 4 Literature Review

The table below summarizes how various data mining techniques have been applied to fashion data to identify trends, categorize garments, make effective decisions, and to modify size charts (Table 1).

Jain and Kumar [1] said that the fashion garment industry has a big amount of data. They used data mining to the fashion industry's production of data i.e., data connected to the variety of garments that is trim, print, fabric, shape, and all to perform garment categorization. They used classification techniques to create a model of classification having 2 categories and sub-categories: first is forecast the garment category, second is forecast the garment sub-category that is upper body, lower body of the garment. Random forest classifier was comparatively much better for the category of garment and sub-category of the garment. Better performance of random forest classifiers it is because it takes a large number of correlated trees which handles unbalanced data in a better way.

Li et al. [3] said that qualitatively evaluate a design effect that uses equipment like EEG and eye-tracking devices for product appearance. How difficult comparative checks the effects of design on product outward form (appearance). Uses some of the equipment like Electroencephalograph (EEG) equipment and eye-tracking devices and eye-tracking device data served as thinking about the activities of human beings.

The study in this paper uses equipment like EEG and eye-tracking devices to record subjects' brain activity and eye-gaze data, and afterwards use some data mining techniques to survey the connection among two different signals. One very important that neural attributes are used to measure out the factors affected by product outward form and rise indicators. We use some research subjects' matter for a case study like women's shirts. Some equipment like the EEG Emotive device and Tobi mobile eye-tracking glasses is useful to track record the subject's brain activity and very important eye-gaze data to measure out the evaluation factors related to product outward form. This method reliably evaluates the particularity of product outward form and provides a purpose reference for better by-product outward form.

**Table 1** Various data mining techniques applied on fashion data [10]

Author Name	Paper Title	Area of data mining	Observation	Year of publishing
Sheenam Jain and Vijay Kumar	Garment categorization using data mining	Garment industry	Categorization of the garment using data mining with two sub-categories	2020
Li Zhao, Chao Min	The Rise of Fashion Informatics: A Case of Data-Mining-Based Social Network Analysis in Fashion	Fashion industry	To monitor customers and understand the fashion world of social media	2018
S. Vinod Kumar and S. Poonkuzhalai	Improvising the Sales of Garments by Forecasting Market Trends using Data Mining Technique	Garment industry	Extracting data to make effective decisions using data mining methods	2018
Sandeep prasad, dr. Vibha lodi assistant professor	A review of statistical and data mining techniques used to prepare standardized size charts	Fashion industry	To avoid pitfalls during the preparation of a size chart	2019
Yanan Yu, Marguerite Moore, Lisa Chapman	Social Media Based, Data- mining Driven Social Network Analysis (SNA) of Printing Technologies in Fashion Industry	Garment industry	To understand the 4- printing technology through the SNA network	2020
Hina Kausher, Sangita Srivastava	Developing Structured Sizing Systems for Manufacturing Ready- Made Garments of Indian Females Using Decision Tree-Based Data Mining	Garment industry	Modifying the international sizing chart of ready-made garments	2019
Zhu-Jun Wang, Cheng Chi, Meng-Yun Zhang, XianYi Zeng, Pascal Bruriaux, JianPing, wang Ying-Mei Xing, Shuo Xu	Data mining-based optimal assignment of apparel size for mass customization	Apparel industry	Sizing selection solution can develop support buying decision	2020

(continued)

**Table 1** (continued)

Author Name	Paper Title	Area of data mining	Observation	Year of publishing
Bao-Rui Li, Yi Wang, Ke- Sheng Wang	A novel method for the evaluation of fashion product design based on data mining	Fashion industry	Qualitatively effects evaluation of the design of product appearance	2017
DwitaAdilah, Andry Alamsyah	An Augmented Method of Selecting Fashion Talent by Adding Social Media Characteristic	Social media industry	We use SVM and random forest methods for examining and predicting social media activity to improve accuracy	2019

Zhao and Min [4] said that the monitor customers and acknowledge the fashion world of social media with the SNA technique of data mining this paper talks about who is an important source for generating good data that is social media which has become a wide and important source for gathering and generating fresh data. Many Fashion companies use techniques of data mining to monitor customer feedback, promote our new collections, reach the target goals of consumers, and know all the comments posted online by people and fashion customers (buyers). some technique is data-mining-based social network analysis is also known as (SNA) technique some information always flows in social media to know about the particular information flow in a fashion show during the Paris Fashion Week and to know more and understand the fashion world on social media. 7 Paris Fashion Week played an important role because it was evaluated in the research to know better how or what data-mining-based SNA will help to recognize the fashion industry. Graph theory and (SNA) converted social networks into models like abstract models of points and lines. One of the social media platforms Twitter was chosen for this study. Twitter is one of the best suitable sources of real-time event content. Through the Twitter (API) Application Programming Interface, data mining technique used, social network analysis by fashion brands to point out a small number of key members of the network to encourage promote time to time new campaigns and adopt the way new outcome.

Poonkuzhalai [5] Said that the data to make an effective decision using data mining methods and help to present forecasting. Extort applicable data from various business units very fast and model them using visual analytics to make very good and effective decisions using the technique of data mining. These analytics helps so many industries like the garment industry to improve the turnover, cost of fluctuating raw material, and good management of balance supply chain. More high growth achieved in the industry can by application some of the technology for clustering and prediction of sales, classification, marketing, production of clothing apparels by means state of art machine learning algorithms. The algorithms of data mining for regression analysis, time series analysis, predictive classification, forecasting, customer preference,

recommendation, or analysis of sentiments, and so on. Have made possible analytics part of the garments industry. A trend analysis and time series analysis sales of any vendors and anyone who are part of any Garment Company are performed in the proposed work. The forecasting sales of Businesses achieved from the existing sales data based on what and which is the customer buying pattern and Customer Relation based reasoning are achieved when we use some of the models like mathematical models on the data. Based on this type of work, the first training and testing of the data with classical data mining algorithms support to show 2 types of terms which is short-term and long-term, forecasting, and evaluation are done on the various or many parameters.

Prasad and Lodi [6] said that the Size chart is very important in the readymade garment industry, rising manufacturing costs resulting from ill-fitting garments, consumer returns. The objective of this paper is to review different data mining techniques to prepare standardized size charts, additive regression, bootstrapping, and predictive mean matching algorithms to impute missing data, multivariate data analysis techniques like Principal Component Analysis (PCA), various clustering techniques viz. K-Means and KMedoid, univariate data analysis for cross-tabulation and size chart computation are used. Finally, the paper concludes with suggestions to avoid pitfalls during the preparation of a size chart. A2 steps clustering approach is put into the chief components: Agglomerative Hierarchical Clustering is also known as (AHC) is used to set the number of clusters. Agglomerative Hierarchical Clustering is a bottom-up approach monitoring the entire starts in its cluster and a set of clusters are combined as one moves up the hierarchy, second step, a non-hierarchical clustering approach, in which the K-means algorithm is used to a bit or we can say a segment the heterogeneous population in deciding the optimum number of clusters.

Yu et al. [7] said that a data-mining drove (SNA) application that is Social Network Analysis for generating a model of the 4 chief printing terms: screen printing, heat transfer, sublimation, and digital printing. SNA uses calculation techniques to record, examine, large social media datasets. This study tells four printing technology through the SNA. Every printing technology shows a group of the corresponding cluster. We use a different color to show the community. Screen printing represents the blue community, tells the features like t-shirt, fashion, clothing tells the apparel decoration techniques in the world of fashion. Digital printing represents the pink community; features that are related to Direct to garment (DTG) printing mention the digital printing directly to the garments heat transfer technologies represent yellow and green communities. Sublimation represents the purple community is a process of digital dye used for polyester fabric or very hard or rough surfaces some examples are a swimsuit, banners, mugs.

Kausher and Srivastava [8] said that the Garment companies are manufacturing companies all garments manufacturing companies are using their design size tables by modifying international sizing charts of ready-made garments. The reason for this study is to tabulate the data which we cover from the variety of figures include height and girth. Mostly 3000 data have been collected with (data mining method survey) Emprise across the ages between 16 and 80 years females to produce sizing

method or structure suitable for clothing manufacture and retailing. The factor analysis techniques are used to control body dimensions to filter from a big amount of variables. Principal Component Analysis (PCA) Method Control body dimensions were selected by reducing the other variables using the PCA technique. Selected key dimensions were used to classify the sample population Decision Tree Method Decision tree algorithm was used for classifying the data through selected key dimensions by PCA technique. One method of data mining called decision tree-based is used to cluster the data.

Wang et al. [9] said that the size solutions for the apparel size regarding the mass customization made to order have been discussed based on some of data mining techniques some of are: K-means clustering, K-means algorithm approaching principle, is a most common and popular partitioning outlook and support vector machine also known as (SVM). SVM was introduced to assign garment size in this work. It is applied in solving pattern recognition, regression problems, and classification. The SVM is a low-cost solution and very easy to understand. 2793 fit models were bound and intent dealing with the standard sizing system. After that, the garment sizes of the target population were assigned using the primary focus, K-means, and SVM methods. Then the solutions were rise by the entire loss of fit, the list of poor fits, and the figure of ideal fit. A sizing selection outlook or point of view can be developed to support their buying decisions. It is important to apply artificial intelligence techniques to improve the correctness accuracy and efficiency due to the complex or knotty relationship between human body types and sizing combinations in future work.

Adilah and Alamsyah [3] said that the expression of fashion modeling and social network open chances for talent to “break in” or “get discovered”. The previous investigation has given out forecast success found on social media presence. We study social network activity to forecast the successful fashion model. We inspect forecasting using classification tasks by making use of Random Forest and (SVM) Support Vector Machine. Investigation finds that social network improves the benefit vary widely and increasing up performed by Random Forest. Support Vector Machines are also known as (SVM) is solve problems like non-linearity, high dimension, and local minimization. We use the Random Forest method in many prediction forecasting problems for its good prediction forecasting ability such as to predict the next color fabric for the next season.

## 5 Conclusion

Data mining with its tool and techniques helps to solve various business problems. The garment industry is comparatively fresh to the department of data mining and machine learning too. In this paper, we reviewed what all work has been done in the field of data mining in the fashion industry. We found out how data mining can be applied to various fashion data to predict customers' purchases or to change size charts or study users' buying behavior. Businesses such as Mynta, Stitch Fix,

Zalando are attempting to get into the future of data to get a better understanding and perception of their consumer bases. Customers' past purchases information is stored in consumer databases which can be mined to make important business decisions regarding consumer purchase behavior.

## References

1. S. Jain, V. Kumar, Garment categorization using data mining techniques. *Symmetry* **12**(6), 2–20 (2020)
2. <https://www.questionpro.com/blog/wp-content/uploads/2019/06/DataCollection.jpg>
3. B.R. Li, Y. Wang, K.S. Wang, A novel method for the evaluation of fashion product design based on data mining. *Adv. Manuf.* **5**(4), 370–376 (2017)
4. L. Zhao, C. Min, The rise of fashion informatics: A case of data-mining-based social network analysis in fashion. *Cloth. Text. Res. J.* **37**(2), 87–102 (2019)
5. S.V.K.S. Poonkuzhali, Improvising the sales of garments by forecasting market trends using data mining techniques. *Int. J. Pure Appl. Math.* **119**(7), 797–805 (2018)
6. S. Prasad, D.V. Lodi, A review of statistical and data mining techniques used to prepare standardized size charts, 581–594 (2019)
7. Y. Yu, M. Moore, L. Parillo-Chapman, Social media based, data-mining driven social network analysis (SNA) of printing technologies in the fashion industry, in *International Textile and Apparel Association Annual Conference Proceedings*, vol. 77, no. 1 (Iowa State University Digital Press, 2020), pp. 1–3
8. H. Kausher, S. Srivastava, Developing structured sizing systems for manufacturing ready-made garments of Indian females using decision tree-based data mining. *Int. J. Mater. Text. Eng.* **13**(12), 571–575 (2019)
9. Z. Wang, C. Chi, M. Zhang, X. Zeng, P. Bruniaux, J. Wang, Y. Xing, Data mining-based optimal assignment of apparel size for mass customization. *Commun. Develop Assembling Text. Prod.* **1**(1), 20–29 (2020)
10. D. Adilah, A. Alamsyah, An augmented method of selecting fashion talent by adding social media characteristic, in *2019 International Conference on ICT for Smart Society (ICISS)* (Bandung, Indonesia, 2019)

# Current State of Engineering Education in India: Student Teacher Scenario in Online Mode



Aniruddh Sharan Mangla, Madhulika Bhatia, Shaveta Bhatia, and Praveen Kumar

**Abstract** The international expansion of abrupt face-to-face training via online media has been necessitated by the emergence of the COVID-19 pandemic. Reacting quickly and thoughtfully to prevent content-based and hand-crafted and design-based engineering from being harmed by the pandemic is imperative. We performed a case study at Amity University, Noida, Uttar Pradesh to assess the state online engineering Studies. For the 2020–21 academic year, 52 faculty members and academic department members 60 students from various departments in engineering groups worked on in the conducted research and provided feedback on their experiences with online teaching in order to help raise awareness of the difficulties they encounter. Our findings indicate numerous obstacles affecting online engineering education including issues with equipment and technological difficulties, learning and teaching challenges, anxiety and relationship issues and physical safety. To provide another example, a large percentage of students demonstrated low interest in the classroom, difficulties staying focused, and expressed mental weariness after having taken several online classes. A review of many studies revealed that online exams were linked to an increase in the teachers' perceptions of cheating. It is critical these findings were derived from this study be shared with other instructors to support effective and selective development of online engineering teaching techniques ahead of COVID-19 and following the pandemic.

**Keywords** Covid-19 · Learning scenario · Teaching scenario · Analysis

---

A. S. Mangla (✉) · M. Bhatia · P. Kumar

Department of Computer Science and Engineering, Amity University Noida, Noida, Uttar Pradesh, India

e-mail: [aniruddh.mangla@s.amity.edu](mailto:aniruddh.mangla@s.amity.edu)

P. Kumar

e-mail: [pkumar3@amity.edu](mailto:pkumar3@amity.edu)

S. Bhatia

Faculty of Computer Applications, MRIU, Faridabad MRIIRS, Faridabad, India

## 1 Introduction

High-level problem-solving abilities are critical to engineering. Various pedagogical approaches, as an example, for instance learning, flipped classroom, and project-based learning, have demonstrated usefulness in the advancement of engineering education.

While online schooling has never been like this before completely unknown to educators, the introduction of the Coronavirus pandemic made online Teaching and learning standards are standardized globally and unprecedented. For the UN's Educational, Scientific, and Cultural Organization (UNESCO), pandemic-induced COVID-19 closures and ensuing educational alterations affected more than 1.5 billion pupils (90.1% of all students enrolled globally). These educational institutions are found throughout the world shut down suddenly, which necessitated the quick conversion of face-to-face instruction into an online (or blended/hybrid) model. Because of this, colleges and universities that were largely based on face-to-face lectures had a variety of difficulties while switching to web-based education.

In order to contain or reduce the effect of a pandemic on engineering education, among the most talented vulnerable, poor, and underrepresented pupils in society (including those who have a difficult time meeting their academic responsibilities due to personal obligations, financial burden, and additional employment) must work expeditiously, diligently, and critically. More work is required to ensure that online engineering training courses continue to fulfil the demanding program accreditation standards [1].

At this time, we believe that there has been no comprehensive (quantitative and qualitative) investigation of the problems and affecting online engineering education's determinants prior to the pandemic. To close this gap, following research questions are important:

As universities that were mostly focused on conventional face-to-face instruction pre-COVID, the biggest challenges that influence virtual education during the COVID-19 pandemic?

How will virtual Education in India be changed in the future?

In order to help higher education online providers, succeed, the consortium of online providers has identified the five main characteristics of online higher education: a set of capabilities that can be of assistance flourish in the workplace, academic satisfaction [2], intellectual fulfilment, accessibility, and cost [3]. With these various factors, we devised a survey and went about finding data with pertinent to the faculty and students at private universities. Surveyed has been taken from many faculties and students and obtained a complete picture of the various difficulties that these participants faced with regards to moving from face-to-face teaching to online teaching in Session 2020–21. This article the outcomes of the research of the surveys and offers ideas for better online engineering education. Letting other instructors know about this study might assist facilitate dynamic engineering instruction as the ongoing pandemic develops. Completing an online engineering degree also receives a boost following the outbreak of the virus period since it helps with complete development.

## 2 Related Work

Our study evaluates the consequences of online education for engineering studies by investigating how students and faculty from different engineering sub-disciplines interact on the internet. Because of this, the work featured in this presentation includes multiple variant traits that set it apart [4]. Investigation uses strong evidence to highlight various strategies we've put forth to assist those institutions that lack money or people, or who serve the community of the first-generation of students [7].

### 2.1 Literature Review

Serial No	Title	Methodology	Limitation
1	Online Engineering Education: Learning anywhere, anytime [5]	Many challenges have been analysed	Some irrelevant challenges found
2	An observational study of engineering online education during the COVID-19 pandemic [6]	Conducted a survey	Some questions were there which would not help in any aspect
3	Online learning integrity approaches: Current practices and future solutions. Online Learning 21 [7]	Suggesting the ideas to faculty	Some approaches were not practical
4	Blended learning: A dangerous idea? The Internet and Higher Education [8]	Adverse effects of hybrid learning	Many situations are based on overthinking
5	A constructivist approach to online training for online teachers [9]	Training faculty to teach with maximum ease	Some training ideas can be a burden on Teachers
6	Virtual laboratories in engineering education: The simulation lab and remote lab [10]	Discussed the trends and issues in virtual laboratories—simulation via the Internet	Only a few issues were discussed
7	A systematic review of research on the flipped learning method in engineering education [11]	Reviewing many articles on flipped learning and analyzing them	The literature used for review was very old
8	Faculty perceived barriers of online Education [12]	Discussion about different barriers of Online education from faculty's point of view	Generalizability of the results

### 3 Methodology and Materials

#### 3.1 Surveys

As we attempted to discover and analyze the distinct issues that our professors and our students faced while taking part in the online teaching session in 2020–21, we created a long-term strategy for online education in the upcoming semesters. The questionnaires were developed holistically, including all input that stakeholders provided to the online teaching, both teachers and students [13]. ‘While faculty survey had 13 multiple choice questions, the student survey consisted of 12 multiple choice questions with extra choice choices for all of the questions.

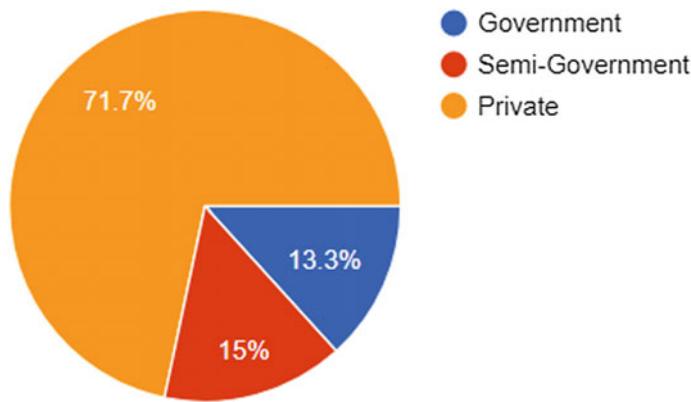
The questions in the faculty survey examined the difficulties that are confronted by professors while using technologies like computers, tablets, styluses, scanners, microphones, and cameras and the reliability of the internet connection. Open-book, closed-book, synchronous, asynchronous examinations, completely online exam (Quiz), semi-online exams (where students solve problems on a paper, then see if the answer has been scanned/upload their solutions). There were questions about proctoring tests and if the teacher says it’s okay believed that cheating was common. Faculty were also asked to specify the areas of their online teaching that they wanted to develop their teaching skills. While participating in online classes, some of the professors were queried about their workload, stress levels, and their engagement with students.

The purpose of these student surveys is to investigate how hard it will be students face while studying online, such as equipment and software not being available, the lack of a stable and genuine internet connection, a quiet/secluded learning environment, and potential conflict with their job or family responsibilities. As well as challenges the synergy that they shared with equivalent online classes (e.g., the absence of teacher focus on technology, or a lack of teacher knowledge about technology) (e.g., Management of time, issues with proctoring photographic approaches employing a camera. These surveys are for the readers’ further reference located within File.

### 4 Survey Results

The poll was completed online over a one-week period using Google Forms. As with the student survey, the conduct of the online survey (Poll) were also all completed online through Google Forms. The survey participants were completely anonymous.

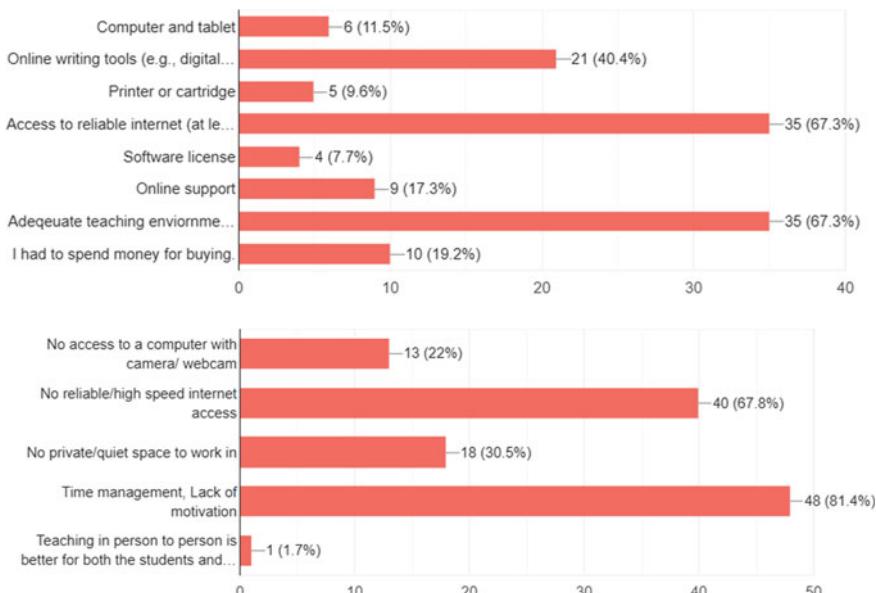
71.7% of respondents were from private universities, 13.3% were from government universities, and 15% were from semi-government universities (Fig. 1).



**Fig. 1** Type of universities

#### 4.1 *Obstacles Faced by Students and Teachers Logistically*

Figure 2 illustrates the different logistical obstacles that survey respondents reported having when completing the online teaching session. Eleven an estimated three-quarters of the faculty members reported problems with the licensing of software or had difficulty using personal computers or tablets. Nearly two-thirds most of the teachers were absent have a consistent internet connection. About 40% the teachers

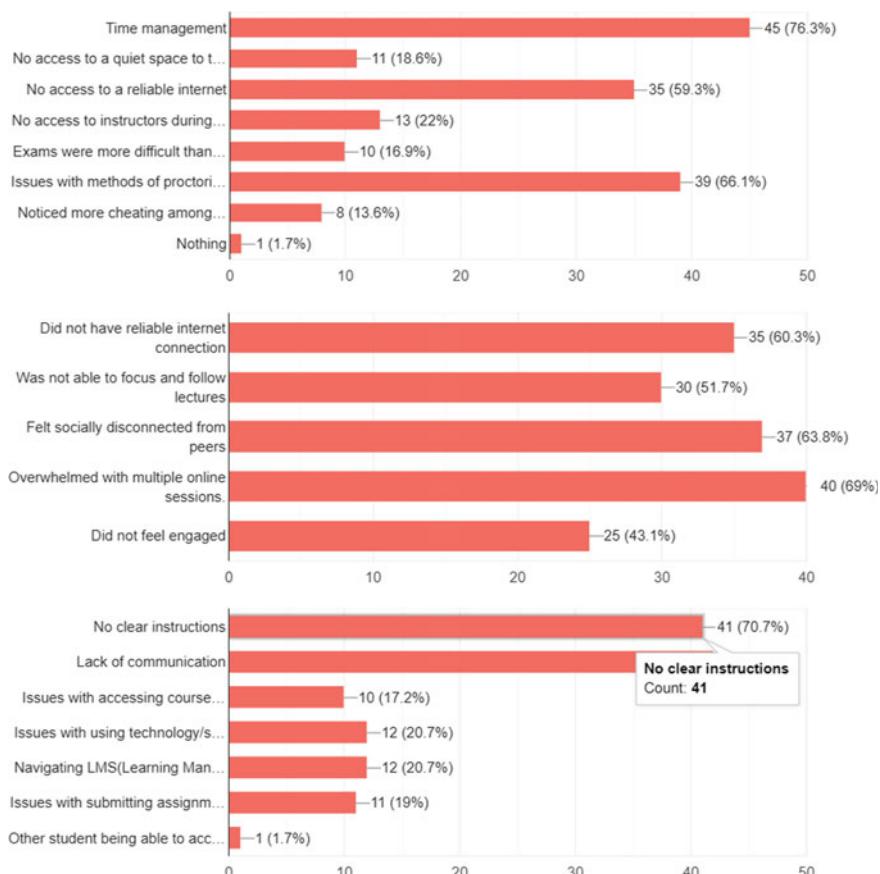


**Fig. 2** Problems faced by students and teachers

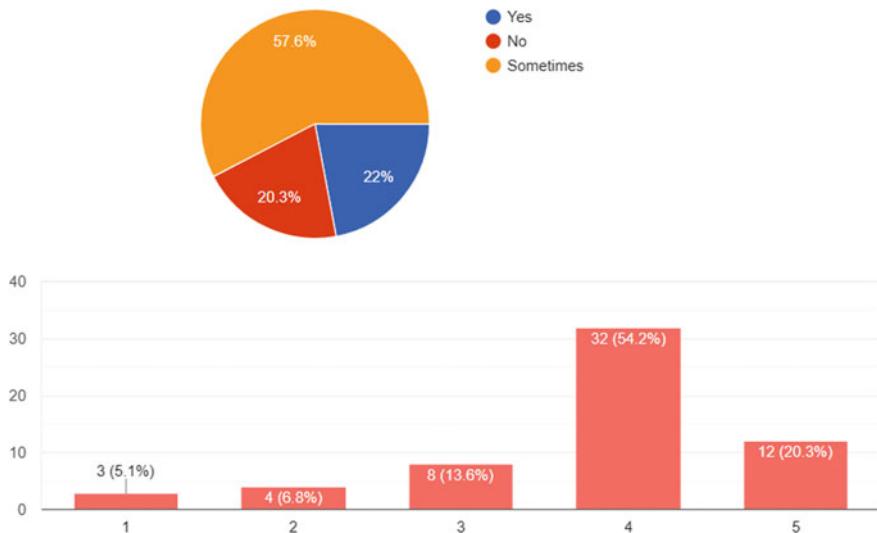
indicated that they had not have online writing tools available to them, or they had difficulty using these resources. Twenty-two percent of the student no access to any computer or tablet, while 68% among the pupils reported that their internet is unreliable, 67% nearly all among the list of pupils selected that they didn't have adequate quiet workspace, and 81 percent of a number of students observed that motivation was lacking and problems with time.

## 4.2 Students' Problems with Online Classes

In Fig. 3, students' problems with online education are summarized. About half of the kids struggled to keep their attention. 70% of pass attendees state that are too tired to do anything afterwards. 64% of students felt disconnected Students with



**Fig. 3** Problems faced by students



**Fig. 4** Stress levels of students

similar interests are peers, while 43.1% of the population didn't feel connected in class while using online learning platforms. 70% the kids were of the belief that professors provided confusing or incomplete direction or communication. Pupils who remain after excluding the failed students have high levels of creativity. Faced issues with uploading assignments and tests online, while 76.3% of students reported time management concerns during online exams. In further comments, students complained about how difficult it was to have their tests run through AI proctor mode, with 33% of students noting difficulties and 66% of students reporting challenges with test methods, such as AI Proctoring. Stress level of students are further shown in Fig. 3.

The other half of the pupils want to hide their camera/microphone. In addition to providing basic personal information, participants also reported having privacy issues with usage of the camera/microphone, especially if they were living in a crowded household or living area shared with others. Furthermore, several students reported an elevated degree of worry caused by the presence of cameras on the computer that impeded their concentration and diminished their exam results. Thirty percent of students are unable to access secluded or quiet spaces in which to do online coursework or to study. Over 80% of pupils didn't have the necessary drive to study. By leaving out the mandatory remarks, the optional comments serve to highlight the difficulty in motivating people in today's workplace: uncertainty about the Corona outbreak pandemic an increased risk of social isolation were shown to be significant contributing causes. Finally, around 15% of the students evaluated their entire online learning experience as positive, with another 30% of those who evaluated it to be negative and the remainder (55%) Some felt neutral about it. People had stressed and anxious feelings when attending online classes, and those numbers

indicate that about 57% of students felt tension and anxiety when using this method. Additionally, approximately 54% of students had significant levels of stress when using this method.

### ***4.3 Type of Assessment Methods Being Used During Teaching in Virtual Mode***

This figure displays a range of assessment methods that professors employ in online classrooms. A exam in which students complete questions on paper, and in this manner, scan and upload solutions is known as a “live semi-online” test. Take-home tests are asynchronous, whereas scheduled classes and tests are synchronized. This research let respondents pick more than one technique for each question, therefore the proportion could not be equal to 100 (Fig. 5).

Around 86% of people observe this of teachers use completely online exams. The savvy consumer has the availability of creating a completely random order and/or querying values with little to no effort. Students can have only one question at a time, and they cannot return to prior questions. Since the efficacy of these alternatives in preventing cheating, and consequently the reduced requirement for future performance, might lower the chance of this technique of testing among intellectuals, there is a chance that this strategy will remain popular may depend on it.

Oral presentations and tests (15%), together with asynchronous semi-online tests (15%). 61%, additionally, our investigation showed that 17% of the unit utilised open book/open tests while 69% used closed book/closed tests. Among geniuses, the requirement for printing equipment can be decreased through the use of open-book/open-note exams. 14% of the group utilised specialised software to perform an action. 80% of the test takers used the Zoom camera and microphone and 20% used some programme to accomplish anything. In addition, we looked into several assessment techniques, and how teachers perceive cheating and conducting an affair those methods. A study done on faculty found that 70% of faculty believe that cheating is gaining ground and that it is primarily. To do with students using computerised testing models like simulations, asynchronous testing, and other methods. The study

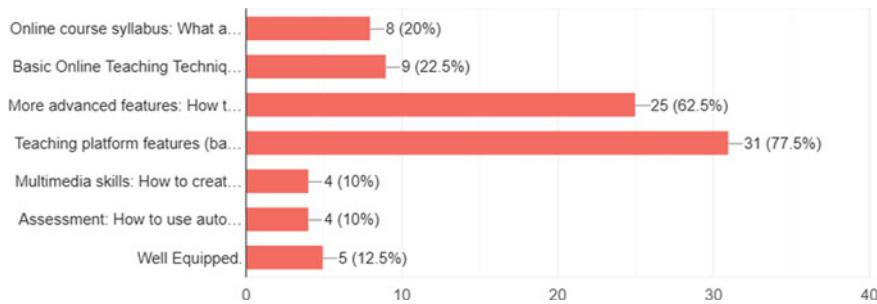
Completely online exams (e.g., MCQ Quiz)

Asynchronous semi-online exam (download, pen and paper, scan, upload)

Synchronous (live) semi-online exam (download, pen and paper, scan, upload)

Oral presentation or demonstration

**Fig. 5** Types of exams



**Fig. 6** Faculty skills that needed enhancement

of the data revealed that the semi-online asynchronous testing, and book-books were associated with an increase in visual fraud.

#### **4.4 Faculty Skills Required to be Enhanced**

Approximately 62% of the faculty will be required to understand the more sophisticated aspects (e.g., the best way to make online surveys, quizzes). Almost 20% of the faculty will need to acquire knowledge on how to construct an the curriculum for an online class or get more proficiency utilizing online learning platforms. In total, 10% as far as the participants were concerned that they would like to enhance their multimedia abilities. More then 50% of the teachers need training in fundamental online learning platform capabilities in order to feel comfortable using them (Fig. 6).

### **5 Discussion**

#### **5.1 Student Challenges**

As a result of the impact of the COVID-19 pandemic, over a quarter of our pupils were not able to connect to a good internet connection, increasing the digital equity gap between them. Coverage of learning outcomes is greatly affected by the ability to access stable internet connection and a personal computer or tablet. While acknowledging public disruptions and frequent site cleaning, the colleges might give WiFi access to campus open areas and well-ventilated facilities to better handle this issue. Loan systems for computer equipment are suited for college students require the technology but can't afford to buy it outright. With the support of academic institutions, college students are have a virtual desktop where they can access all necessary applications. Smartphones and tablets that come equipped with free scanning applications can help rectify the lack of availability of scanners.

Peer support and communication difficulties were found in online also, classes, with a lack of attention, focus, and clear instruction from instructors. Additionally, they found themselves at a disadvantage with regard to time management. The finding is that peer support and cooperation help increase student success levels, specifically for marginalized student populations groups. When online education was a new concept during the COVID-19, students suffered from a lack of peer support. Possible solutions to these issues can be partially resolved by using faculty teaching strategies designed to enhance student learning by shortening lessons into portions where the text is kept short breaks, encouraging students to work together in groups, providing enthusiastic and interested pupils effective online course, and helping students record lectures. Live broadcasts that are still going on when the dialogue has concluded. Students who are struggling to study should be aided by the latter. Faculty may help to alleviate the issue of time management during exams by creating sample tests to help students adjust their test format ahead of time.

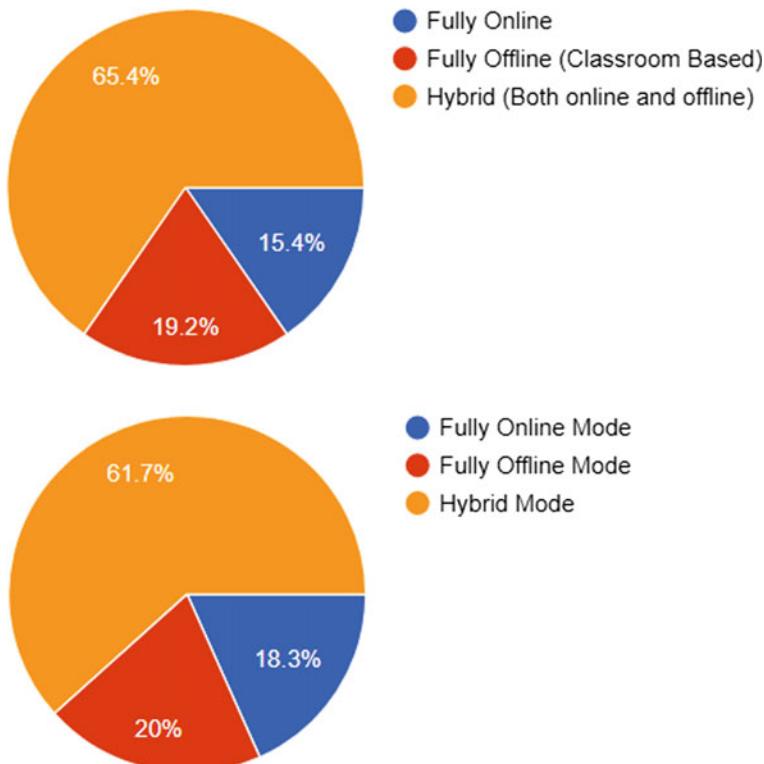
## **5.2 *Challenges Faced by Teachers***

This is essential to online education because institutions create quality standards for online education. A large percentage of our participants stated a lack of access to the relevant software, hardware, and related tools. Many clever people have mentioned that they don't have an online writing tool without a real white board in the classroom. Allocating an institution's money for Hardware and Tools would solve this difficulty.

## **5.3 *Prediction of the Future of Online Engineering Education in India***

Additionally, we also wanted to know students and faculty as to what type of teaching mode they would prefer one day. Our survey revealed that, when it comes to students, 61% would prefer the hybrid method of teaching, which combines both on-campus and online courses, and when it comes to faculty, 65% prefer the hybrid method of teaching, which also combines on-campus and online courses.

The study's findings suggest that majority of the participants prefers to attend university-sponsored classes on campus, in addition to studying via the internet. As a result, we can reasonably conclude that few, if any, individuals want complete virtual model. The future of online engineering education in India is thus not very bright (Fig. 7).



**Fig. 7** Choice of Mode of Study

## 6 Conclusion

To accurately assess the obstacles, we experienced during the COVID-19 outbreak due to the preface of online teaching of Whilst studying engineering subjects, we created a research project that measured we and our students technology at our low-performing university. Technical, organizational, and preparation-related problems were discovered, and several treatments were put up to deal When you're among them. This is as a result, we found that the future of online engineering education is not the same bright as many may hope, everyone want to be educated online, but there are a limited number of options. The vast majority of students want to attend college to complete their education while also having the option of taking classes online.

**Acknowledgments** The Author would like to thank all the Participants for participating and sharing their opinion in the surveys.

## References

1. S. Asgari, B. Penzenstadler, A. Monge, D. Richardson, computing to change the world for the better: A research-focused workshop for women (IEEE, Portland, Oregon, USA, 2020)
2. M.A. Almaiah, I.Y. Alyoussef, Analysis of the effect of course design, course content support, course assessment and instructor characteristics on the actual use of E-learning system. *IEEE Access* **7**, 171907–171922 (2019)
3. E.M. Aucejo, J. French, M.P. Ugalde Araya, B. Zafar, The impact of COVID-19 on student experiences and expectations: Evidence from a survey. *J. Public Econ.* **191**: 104271.pmid:32873994 (2020)
4. P. Tilak, M. Deshmukh, S. Phadake, A survey on online examination during COVID 19 pandemic: Perception of Management Students. *Mukt Shabd J.* **9**, 284–290 (2020)
5. H. Memarian, Online engineering education. *Iran. J. Eng. Educ.* **21**(82), 15–29 (2019)
6. H. Suryaman, H. Mubarok, Profile of online learning in building engineering education study program during the COVID-19 pandemic. *IJORER: Int. J. Recent Educ. Res.* **1**(2), 63–77 (2020)
7. A. Krouská, C. Troussas, C. Sgouropoulou, Fuzzy logic for refining the evaluation of learners' performance in online engineering education. *Europ. J. Eng. Technol. Res.* **4**(6), 50–56 (2019)
8. Z. Lei et al., Unified and flexible online experimental framework for control engineering education. *IEEE Transactions on Industrial Electronics* (2021)
9. D. Valiente et al., Active learning program supported by online simulation applet in engineering education. *SIMULTECH* (2019)
10. M. Barr, S.W. Nabir, D. Somerville, *Online Delivery of Intensive Software Engineering Education During the COVID-19 Pandemic*, 2020 IEEE 32nd Conference on Software Engineering Education and Training (CSEE&T) (IEEE, 2020)
11. S. Asgari et al., An observational study of engineering online education during the COVID-19 pandemic. *Plos one* **16**(4), e0250041 (2021)
12. A. Ukkivi et al., Mathematics online learning model in engineering education. *INTED 2019 Proc.* 8941–8947 (2019)
13. B. Hassan, A.A. Shati, A. Alamri, A. Patel, A.A. Asseri et al., Online assessment for the final year medical students during COVID-19 pandemics; the exam quality and students' performance. *Onkologia i Radioterapia* **14**, 1–6 (2020)
14. D. May, Cross reality spaces in engineering education–online laboratories for supporting international student collaboration in merging realities. 4–26 (2020)
15. C.M. Clevenger, et al., Assessing an online tool to promote sustainability competencies in construction engineering education. *J. Prof. Issues Eng. Educ. Pract.* **145**(1), 04018014 (2019)

# Estimating Blockchain Using Time-Series Forecasting ARIMA



Vaibhav Bhatnagar and Bhoomika Batra

**Abstract** The purpose of the study is to forecast the bitcoin prices fluctuation for a month based on the last ten years. The data is secondary and collected from investing.com using the time series technique ARIMA. The findings state there are price fluctuations seen, it is increasing at a low rate. ARIMA model accurately predicts the direction of travel; nevertheless, the predicted values are consistently lower than the actual observations. AI is being used to estimate the value of bitcoins as a means of measuring the potential risk connected with blockchain technology's financial use is being investigated further work that may be followed up based on this research.

**Keywords** ARIMA · Bitcoins · Cryptocurrency · Artificial intelligence (AI) · Blockchain

## 1 Introduction

The digitalization of economies is perhaps the most significant disruption now affecting all economies including financial institutions. Distributed ledger technology whose most well-known use is the cryptocurrency Bitcoin, is the most up-to-date method of creating and consuming digital products. Volatility in the price of nontangible digital products, notably Ethereum, is a major source of anxiety. Cryptocurrencies are the latest era of digital currencies that have piqued the interest of economic experts. It has various characteristics like Cryptocurrencies are extremely secure. Bitcoin's value and number of transactions surged drastically since it first attracted global attention at the end of 2013. The probability of BTC returns has been investigated using several variables such as social media coverage [7]. Furthermore, [5]

---

V. Bhatnagar (✉)

Manipal University Jaipur, Jaipur, India

e-mail: [vaibhav.bhatnagar@jaipur.manipal.edu](mailto:vaibhav.bhatnagar@jaipur.manipal.edu)

B. Batra

V.V. Giri National Labour Institute, Noida, India

looked into the effect of comment threads on social media platforms on virtual currency price volatility and transaction volume, and observed that the quantity of favourable social media comments is strongly linked to BTC. To anticipate price fluctuation using BTC historical data, researchers employed a regression method that employs the latent source model proposed by [3]. Researchers were capable to achieve a 55 percent categorization accuracy.

**Adcock and Gradojevic** [1] employed an ANN to attain a 65% classification accuracy. Furthermore, [6] To predict the BTC price, researchers used a various neural network technique. Using LSTM, author was able to attain a classification accuracy of 52% with an RMSE of 8%. To anticipate the lowest, maximum, and closing BTS prices, researchers used various higher algorithms in regression models, with the SVM approach outperforming the others with a MAPE of 1.58% [4]. With one of the most recent studies on forecasting BTC daily prices, researchers combined high-dimensional features with numerous machine learning algorithms.

## 2 Methodology

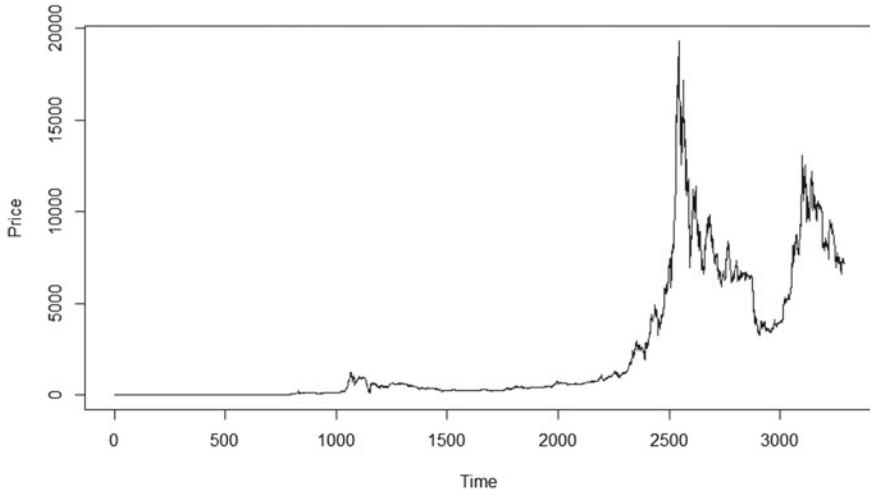
The goal of this research is to use machine learning to forecast BTC prices over time. The underlying interdependencies in time-series forecasting of bitcoin values are difficult to grasp and model. Variance and standard deviation, for example, are statistical parameters that alter over time. The BTC price technical indicators were collected using accessible data sources in our research. Data is collected, cleaned, and scaled/normalized at the data pre-processing stage. The BTC data is sorted and put into three groups.

### 2.1 Dataset

BTC features and price data are publicly available on the internet. The information for this study came from investing.com. The total number of observations for this study is 3288 ( $n = 3288$ ), with the geometrical traits of non-linear trend and non-stationarity shown in Fig. 1, which reflects the raw data. BTC prices are demonstrated to be a static time series using unit root augmented Dickey-Fuller (ADF) testing, with an ADF statistic of 2.748 at the 5% significant level ( $p\text{-value} = 0.29$ ).

### 2.2 Statistical Approach

Time-series forecasting will be implied in the current study using the linear forecasting model ARIMA. A time series is a collection of observations that are measured through time. The observations at various points in the series have varying degrees



**Fig. 1** Series are non-stationary

of correlation with one another. Because the values of successive observations are interdependent, future values can be anticipated based on previous ones. Linear models, such as the ARIMA, are well renowned for their accuracy and flexibility in the prediction of linear time series, as well as some successful short-term value or trend predictions in non-linear processes. It's vital to split the relevant dataset into multiple groups when choosing prediction models: training and test data. ARIMA is the first algorithm we'll employ to attempt to anticipate Bitcoin's price. When the author uses very minimal time, the outcome of ARIMA just isn't very informative because it uses prior data to forecast a value. The Autoregressive Integrated Moving Average (ARIMA) is a linear model that incorporates an AR, an MA, and an integrated component that differentiates the time series to transform it into a stationary series.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t \quad \text{AR} \quad (1)$$

where  $\varepsilon_t$  is white noise,  $-1, -2, -p$  is the lagged values of  $y_t$ .

$$y_t = \alpha + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} \quad \text{MA} \quad (2)$$

where  $t$  stands for white noise, this is known as an MA( $q$ ) model, or a  $q$ -order moving average model. The ARMA model combines the strengths of both AR and MA elements. An ARMA ( $p, q$ ) model includes the  $p$ th order AR and  $q$ th order MA models in this way. The AR and MA coefficients vectors are denoted by  $\mathbf{Y}_t$  and  $\mathbf{y}_t$ . The intercept and error term at time  $t$  are captured by  $\alpha$  and  $t$ . In Eq. 3, the whole ARMA ( $p, q$ ) model can be shown in depth ( $\mathbf{y}'_t$ )

$$\begin{aligned} y_t' = & \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + \epsilon_t \beta_1 \epsilon_{t-1} \\ & + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad \text{ARMA} \end{aligned} \quad (3)$$

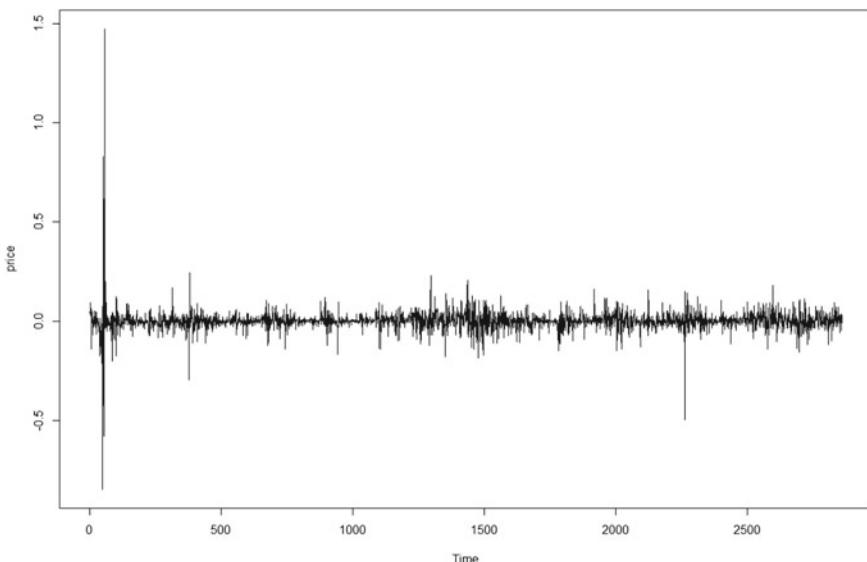
where  $y_t'$  denotes the truncated series (it may have been differenced more than once). Both lagged  $y_t$  values and lagged errors are included in the “predictors” on the right-hand side.

### 3 Data Analysis

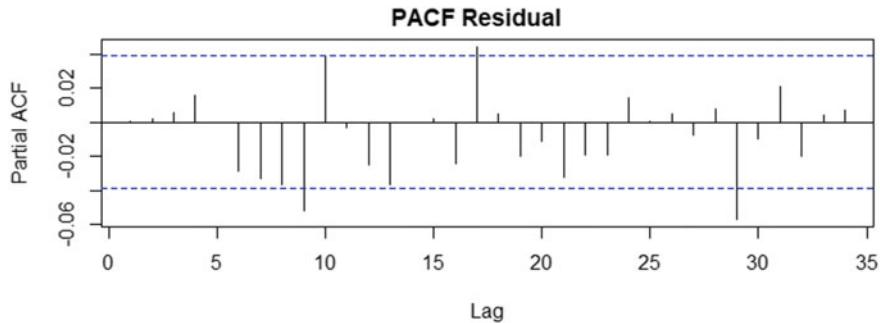
The current study model will be developed using R, a time series that is often used to see what an entity evolves through time. The only assumption ARIMA has to follow is that the series must be stationary. The term “stationary” refers to the fact that the mean, variance, and covariance remain constant across time. Using the difference between each data point though, the series can be stationary. The level of differencing between data points required to render the data stationary is represented by the  $d$  in ARIMA ( $p, d, q$ ). Initially, the series be stationary or non-stationary were checked through the Augmented Dickey-Fuller test (ADF) (Refer Fig. 2).

Then, the author started the identification of Auto-Regressive and Moving Average from the series using it's functions.

The Fig. 3 shows substantial correlations at the first or second lag, followed by nonsignificant correlations. In the series, for the Autoregressive term. The order of



**Fig. 2** Series are stationary



**Fig. 3** Partial autocorrelation function of series



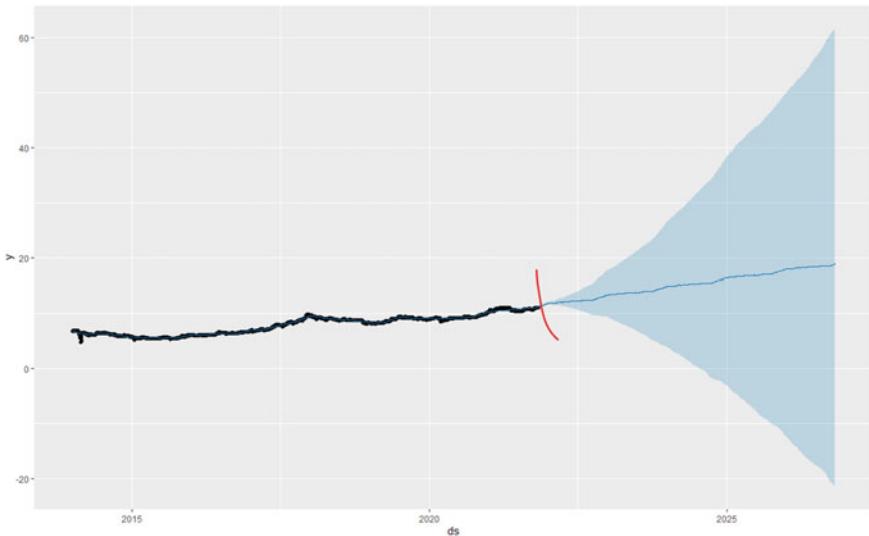
**Fig. 4** Actual bitcoin prices from 02.12.2019 to 01.01.2020

the autoregressive term is determined by the number of significant correlations. To apply **ARIMA**, the ARMA order must be stationary at 0 levels, therefore the data is eligible to use ARIMA [2]. The Akaike information criterion (AIC) is a quantitative technique for estimating how often a model is correctly specified it must have been constructed from. AIC is a mathematical tool for evaluating various models and selecting whichever model fits the data better. It can be used to choose the ARIMA model's order (p, d, q). For the study, the ARIMA model is (1, 0, 1). At a later stage, we forecast the bitcoin prices using the forecast function (Fig. 4).

$$2.21 + \beta_1(-0.7581) + \beta_2(0.8076)$$

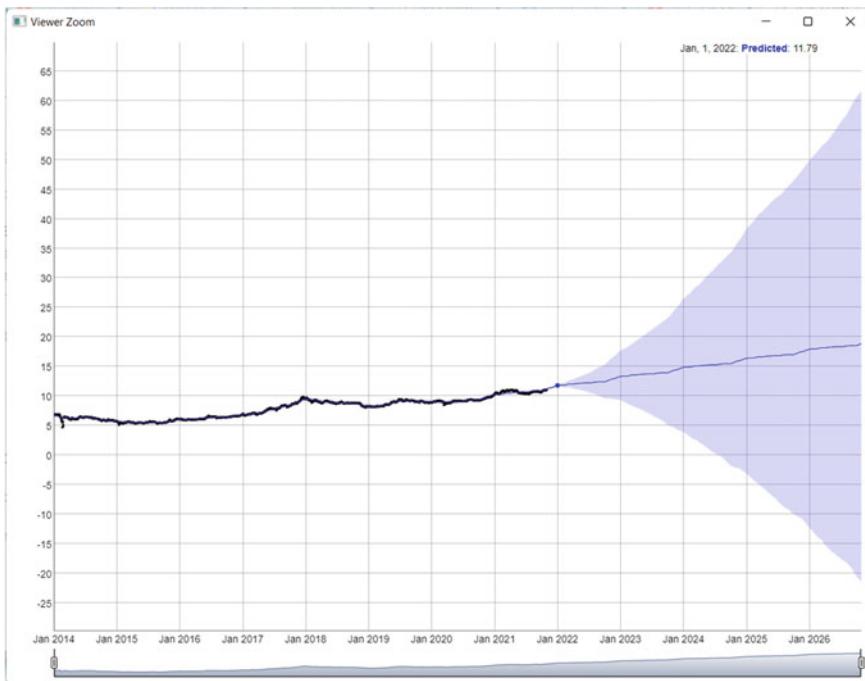
## 4 Discussions and Future Scope

In this article, a forecasting model ARIMA has been used to predict the crypto bitcoin prices taken for 10 years on daily basis and forecasted the next 30 days' prices. In forecasted values from Fig. 5, there are price fluctuations seen, it is increasing at a low rate. Machine learning approaches have gained a lot of traction in the worldwide community recently. The major goal of this dissertation is to forecast and predict



**Fig. 5** Forecasted bitcoin prices from 02.01.2020 to 31.01.2025

crypto bitcoin prices over the next ten years. When it came to the ARIMA model, the best fit was an ARIMA (1, 0, 1) based on the minimum AIC criteria. This model accurately predicts the direction of travel; nevertheless, the predicted values are consistently lower than the actual observations. This approach could potentially be used to spot fraudulent activity and unusual behavior. Additional data inputs linked to global happenings and worldwide financial risk could be employed to examine and anticipate the sustainability of cryptocurrencies utilizing a mixture of a machine attempting to learn price algorithms and anomalous detection techniques. Further, the research could look into the effectiveness of simpler variants of neural networks like ANN to forecast. It would also be fascinating to anticipate bitcoin prices for several days in the future using historical data (Fig. 6).



**Fig. 6** Comprehensive predicted graph

## References

1. R. Adcock, N. Gradojevic, non-fundamental, non-parametric Bitcoin forecasting. *Physica A* **531**(121727), 121727 (2019). <https://doi.org/10.1016/j.physa.2019.121727>
2. R. Bevans, An introduction to the Akaike information criterion, Scribbr. Available at: <https://www.scribbr.com/statistics/akaike-information-criterion> (Accessed: 19 Jan 2022)
3. G.H. Chen, S. Nikolov, D. Shah, A latent source model for nonparametric time series classification, *arXiv [stat.ML]* (2013). Available at: <http://arxiv.org/abs/1302.3639>
4. Z. Chen, C. Li, W. Sun, Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *J. Comput. Appl. Math.* **365**(112395), 112395 (2020). <https://doi.org/10.1016/j.cam.2019.112395>
5. S. McNally, J. Roche, S. Caton, *Predicting the Price of Bitcoin Using Machine Learning*, in 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP) (IEEE, 2018)
6. D. Philippas et al., Media attention and Bitcoin prices. *Financ. Res. Lett.* **30**, 37–43 (2019). <https://doi.org/10.1016/j.frl.2019.03.031>
7. A. Urquhart, What causes the attention of Bitcoin? *Econ. Lett.* **166**, 40–44 (2018). <https://doi.org/10.1016/j.econlet.2018.02.017>

# Analyzing Online Reviews of Customers Using Machine Learning Techniques



Anandhavalli Muniasamy and Roheet Bhatnagar

**Abstract** Web data mining can provide insight into the e-commerce data which will be useful to promote sales, understand customers, and support new business opportunities. But understanding the customers and increasing their participation in online shopping is a great challenge as it requires the analysis of the relationship between customer satisfaction and the most significant factors that influence their online buying decisions. Machine learning comes into play in addressing the challenge faced by the e-commerce companies. In a nutshell, this paper focuses on the analytics of customers' attitude towards using online shopping and intention to buy, using machine learning classifiers namely Naive Bayes, Logistic regression, Support Vector Machine, and Neural Network. For data analysis, Latent semantic analysis is applied to examine the most frequent words used in the online reviews. Finally, customer's interest in online shopping analysis has been performed using machine learning classifiers to classify the customers' sentiment from their posted reviews on the e-commerce platform. Also, we compared the performance results of these classifiers on the ecommerce dataset. The results reveal that the Naive Bayes classifier has performed better than all the other three classifiers in terms of execution time and the measures like accuracy, recall, F1-score, confusion matrix.

**Keywords** Online shopping · Machine learning · E-commerce · Latent semantic analysis (LSA) · Naive bayes (NB) · Support vector machine (SVM) · Logistic regression (LR) · Neural network (NN)

---

A. Muniasamy

College of Computer Science, King Khalid University, Abha, Kingdom of Saudi Arabia

R. Bhatnagar (✉)

Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, Rajasthan, India

e-mail: [roheet.bhatnagar@jaipur.manipal.edu](mailto:roheet.bhatnagar@jaipur.manipal.edu)

## 1 Introduction

Online shopping is a form of e-commerce and provides advantages such as avoiding store visits and travel costs, shopping at any time, increasing marketing, reduction in the stock & inventory related expenses, and have proved to be beneficial for both the retailers and customers. According to Umair [1], it has been found that more than 85% of the world's online population supported the online shopping via internet during the recent year. This involves the technical specifications of an online store focused on the technology-centered view, which influences consumer's awareness of using that technology [2]. It allows the customer to buy products or services directly online by choosing products while visiting online shopping mobile applications or websites [3].

The study [4] sheds light on an important topic in the field of e-marketing, which is accepting online shopping by consumers, and discussed the benefits of using online shopping in terms of saving money, time, effort, and getting rare products at competitive prices, and its impact and the results of using Internet in marketing, promotion, and sales to reduce operational costs.

This study [5] helps in understanding and analyzing the factors affecting individuals' behavior and decisions towards adopting or using internet shopping and ensuring their confidence in using internet shopping technology as the human factor is the main driver of this process.

The paper aims to apply four classifiers namely Neural Network, Support Vector Machine, Naive Bayes and Logistic Regression on e-commerce dataset, and compare the performance of the classifiers using various metrics in mining the buying interest of the customers in online shopping. The remaining paper is organised as follows: Sect. 2 focuses on survey related to online shopping and customer interests in e-commerce. Section 3 experimental methodology. Section 4 covers the result and analysis, and the conclusion of the paper is summarized in Sect. 5.

## 2 Literature Review

According to Gooding et al. [6], the number of online shoppers in US during 2019 has reached 93 million, as an increase of 35 million compared the year 2018. Authors in [7] reported that more than 60% of consumers shopping online at least once a month and only 1% have never shopped online. Thus, competition between online shopping sites has become intense due to the profitable marketing through the Internet. Factors like the basic way to get the product and the store's environment may affect the evaluation and selection of the product by the customer and thus environmental factors are very important to enhance the shopping behavior as the study showed [8]. The study [9] analyzed the online customer behavior, and reported that the websites social features, compatible with lifestyle and online customer services has a significant impact on customer behavior and their intentions to adopt online purchase

services, the number of visits by other consumers to the product had a positive impact on the dimensions of financial situations and compulsive buying online [10]. e-commerce is the transaction channels in which people, organizations, and governments all participating for doing business transaction [11]. Despite the several advantages of e-commerce, the study [12] shows that there are some technical problems that e-commerce faces, such as the lack of understanding of the users' interests, due to the limited interaction between human and computer. In addition to the fact that using the same keywords in searching across different platforms may lead to unsecured results due to huge data and also lack of customization. The study [13] also discussed the problem of inflexible interactive mode like text input and mouse click on ecommerce transaction rather than the voice interaction which is more convenient.

Extracting useful information from complex big data is a hot research topic in the recent years due to the application of various classification algorithms for improving ecommerce through the use of web data mining, feature extraction concept, text categorization and clustering, web text association analysis etc. for accurately extracting the text data in the web pages [14]. In online shopping, the product page has an important role in attracting customers to the portal [15].

LSA is an algebraic-statistical method [16] while, LR is a supervised learning algorithm and is a form of the extended regression model to classify data and attain results of two or more possible values. LR is useful to perform sentiment analysis based on text classification and to examine the relationship between variables to investigate the most influencing factors on users' buying decision [17].

NN have successfully been used in various classification problems and recently it has been used for text mining. Neural networks techniques are used to increase the accuracy of classifiers learned with standard machine learning algorithms by obtaining better vector representations for words [18]. NB is a straightforward probabilistic classifier that uses Bayes' theorem and strong independence assumptions to classify the given data. Independent feature model is a more descriptive name for the basic probability model. Because of these feature independence assumptions, the sequence of features is unimportant, and the presence of one feature does not influence the presence of other features in classification tasks [19] while using NB.

SVMs requires little training data to begin producing correct results, but on the other hand, requires more computing power than Naive Bayes, and the outcomes are quicker and more accurate [20]. Each of the reviewed works showed different factors affecting the adoption of online shopping and the buying decision, and many studies presents the trust as an important factor for customer behavior. Therefore, this research proposed using machine learning techniques to investigate the customer's interest towards the adoption of online shopping and focuses on the impact of trust.

**Table 1** Description of datasets

Name of the attribute	Data type	Description
Clothing ID	Categorical	Specific piece being reviewed
Age	Integer	Reviewer's age
Title	String	Title of the review
Review text	String	Review body/text
Rating	Ordinal integer	Product score from 1 Worst to 5 Best
Recommended IND	Binary variable	Customer recommended –1 and 0—Not recommended
Positive feedback count	Integer	Documenting the number of other customers who found this review positive
Division name	Categorical	Name of the product high level division
Department name	Categorical	Name of the product department name
Class name	Categorical	Name of the product class name

### 3 Experimental Methodology

#### 3.1 Dataset

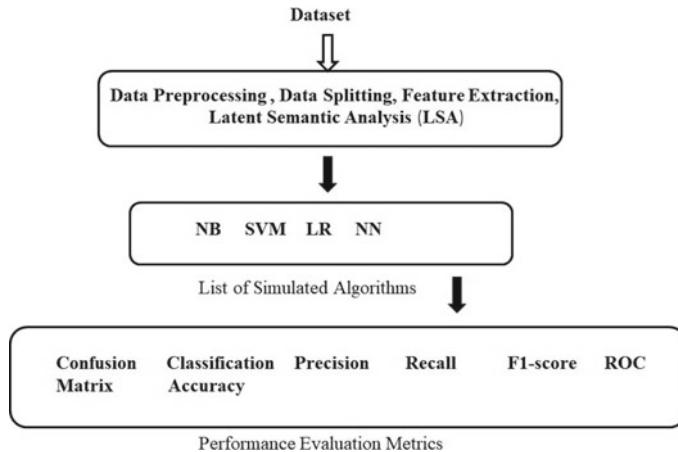
For the experiments, we used Women's Clothing E-Commerce dataset [13] shown in Table 1. This dataset has 10 feature variables and 23,486 rows. Each row represents a customer review, and the description of the data features are given in Table 1.

The column (Recommended IND) is used to represents the sentiments, which is a binary variable stating 1 for product recommended, 0 for not recommended.

#### 3.2 Methodology

The experimental methodology used in this paper is shown in Fig. 1. Firstly, we preprocessed the dataset before simulating the classification algorithms. All the experiments are performed in Python platform. We used the ratio of 80/20 data split for training and testing sets. For the classification of the customer recommendation for the product and their review, the four classifiers are applied to the preprocessed datasets and the performance of every classifier is analyzed using the metrics given in Fig. 1.

Online customer reviews are text data which is mainly consisting of nouns, adjectives, adverbs, verbs, and emoticons. Before applying customer's review analysis, the data should be prepared in a form that machine can understand and deal with. The ecommerce dataset is pre-processed using several NLP techniques to prepare data for analysis. Then, data is analyzed using LR, SVM, NB, and NN classifiers to implement sentiment analysis.



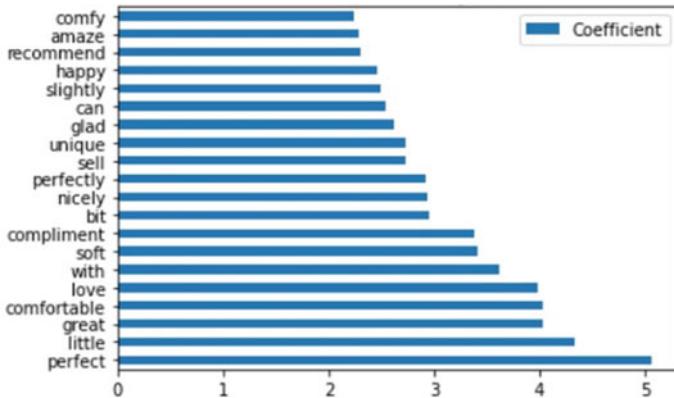
**Fig. 1** Experimental methodology

### 3.3 Data Preprocessing and Analysis

Natural language processing (NLP) tools were imported from NLTK library. The main NLTK tools used for pre-processing for data analysis, firstly, Latent semantic analysis (LSA) is applied to examine the most frequent words used in the online reviews. Then a word cloud is created for clearer view, Finally, a sentiment analysis is performed using four main classifiers to classify the customers' sentiment from their posted reviews on the e-commerce platform. LSA has been applied to know the top 20 frequent used words and their weight are as follows Figs. 2 and 3.

**Fig. 2** Top 20 words with coefficients in recommended reviews

	Coefficient
perfect	5.059418
little	4.332028
great	4.038365
comfortable	4.036251
love	3.983114
	Coefficient
unflattering	-4.005775
cheap	-4.324764
be	-4.384440
disappointed	-5.105635
return	-5.121882



**Fig. 3** Top 20 words in recommended reviews

### 3.4 Feature Selection

In order to reduce the computational cost of modeling and/ or to increase the model's accuracy, the number of input variables should be reduced and Feature Selection is used. We concentrated on two features for our task: the review text and the Recommended IND. The “review text” includes all the online reviews from customers, it is selected to be analyzed and mining the customer interest using text mining techniques.

### 3.5 Machine Learning Classifiers

Four classifiers are trained using the training data, which were then applied on to the test data to predict users' sentiments. The column (Recommended IND) is used to represents the sentiments, “Recommended IND” includes Binary variables stating where the customer recommends the product, 1 is recommended, 0 is not recommended. We consider 0 for negative sentiment and 1 for Positive sentiment. Finally, different performance metrics are used to evaluate and compare between the four classifiers.

## 4 Results and Analysis

ML Classifiers were analyzed on the basis of six performance metrics. Table 2 depicts the description of all the metrics used in this study.

**Table 2** Various matrices

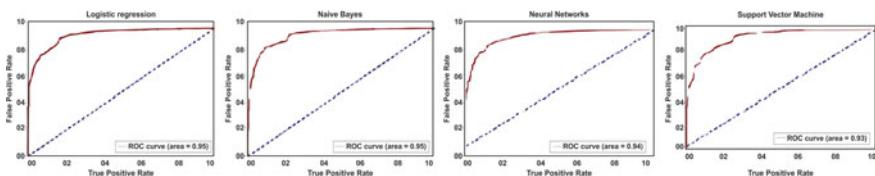
Name of the metrics	Formula		
Confusion matrix	Actual	Predicted	
		0	1
		True Negatives	False Positives
	1	False Negatives	True Positives
Classification accuracy	$\frac{\text{True Positives} + \text{True Negatives}}{\text{Total}}$		
Precision	$\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$		
Recall	$\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$		
F1-Score	$2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$		

Evaluation with the receiver operating characteristic curve (ROC) are shown in Fig. 4, which shows that results look appropriate for all the classifiers.

Confusion matrix (CM) measure used while solving classification problems. It shows the classification measures of the given dataset in a specific table layout that allows visualization of the performance of a classifier. It has been found that SVM classifier does not give good results although it has high ROC values.

Mere examination using ROC curves and Confusion Matrix are not good enough for a final decision, thus a classification report is presented in a tabular format as part of Table 3. It enlists all the four classifiers results based on the popular standard performance matrices.

LR and NB achieves the best F1-score values among other classifiers (0.96 and 0.97). To decide which one is the best from these two classifiers, the elapsed time in the classification process is calculated for each classifier, NB consumed the least

**Fig. 4** ROC for each classifier**Table 3** Comparison between the classifiers' results

Classifier	Precision	Recall	F1-Score	Accuracy	Execution time
LR	0.87	0.94	0.96	0.94	0:00:01.143353
NB	0.96	0.98	0.97	0.94	0:00:00.012192
SVM	0.79	0.89	0.94	0.89	0:00:58.608685
NN	0.96	0.97	0.96	0.93	0:01:59.735018

execution time. Therefore, NB is considered as the best classifier in this research work.

## 5 Conclusion

The e-commerce domain is very vibrant in modern times, and generate lot of data pertaining to online transactions which can be easily handled & managed by advancement in ICT technologies. Finding the hidden knowledge in this data is a challenging task in machine learning. The current work explores different well-known machine learning algorithms for the analysis of customers interest on online shopping and classify the customers' sentiment from their posted reviews on the e-commerce platform. Also, the authors have compared the performance results of these classifiers on a standard e-commerce dataset. The results reveal that Naive Bayes classifier has performed better than all the other three classifiers LR, SVM and NN in terms of execution time and the metrics like 'accuracy, recall, F1-score, confusion matrix and ROC'.

## References

1. C.M. Umair, The trend of online shopping in 21st Century: impact of enjoyment in TAM model. *Asian J. Empirical Res.* 131–141 (2013)
2. C.H. Cheng, A study on the applications of data mining techniques to enhance customer lifetime value. *WSEAS Trans. Inf. Sci. Appl.* 319328 (2009)
3. M. Kütz, Introduction to E-commerce. (2016)
4. B. Mohammed, M. Mouhoub, Evaluation of an online shopping system under preferences and constraints, in *Canadian Conference on Electrical and Computer Engineering* (IEEE, 2014), pp. 1–8
5. K. Moore, Ecommerce 101+ the history of online shopping: what the past says about tomorrow's retail challenges. Retrieved from big commerce: <https://www.bigcommerce.com/blog/ecommerce/#faqs-about-ecommerce>. (2020)
6. A. Gooding, et al., TOSNOS: to online shop, or not-to online shop-enabling combined improvements, in *2020 IEEE International Conference on Pervasive Computing and Communications Workshops* (IEEE, 2020, March), pp. 1–3
7. A. Nanji, Online shopping trends 2013: most popular categories, top purchase drivers. (2013)
8. G. Mohan, B. Sivakumaran, P. Sharma, Impact of store environment on impulse buying behavior. *Eur. J. Mark.* (2013)
9. H.H. Cheng, T.J. Fu, The Determinants of online shopping behavior, in *2018 International Conference on Intelligent Autonomous Systems (ICoIAS)* (2018), pp. 97–100
10. N.A. Napitupulu, Z. Hidayat, The influence of online shopping applications, strategic promotions, and hedonist habits on e-shopaholic behavior, in *2020 International Conference on Information Management and Technology (ICIMTech)* (IEEE, 2020, August), pp. 922–927
11. K. Johan, W. Samantha, M.J. Tandean, S.O. Sihombing, The relationships between web design, reliability, privacy, service quality, and purchase intention of customers at e-commerce business: an empirical study. (2020)
12. Y. Huang, Y. Chai, Y. Liu, J. Shen, Architecture of next-generation e-commerce platform. *Tsinghua Sci. Technol.* 24(1), 18–29 (2018)

13. Nicapotato. Women's e-commerce clothing reviews. <https://www.kaggle.com/nicapotato/women-e-commerce-clothing-reviews>. (2018)
14. W. Lin, L. Dongying, S. Haizhang, D. Shengbao, Research on the evolution law of the semantic web structure of online shopping reviews, in *2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (2017), pp. 395–398
15. K.M. Sam, C.R. Chatwin, Evaluating the effectiveness of online product planning and layout tools in online apparel shopping, in *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 635–639 (2015)
16. N. Evangelopoulos, Citing Taylor: tracing Taylorism's technical and sociotechnical duality through latent semantic analysis. *J. Bus. Manage.* 57–74 (2001)
17. W.P. Ramadhan, S.A. Novianty, S.C. Setianingsih, Sentiment analysis using multinomial logistic regression, in *International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)* (IEEE, 2017), pp. 46–49
18. N. Garg, K. Sharma, Machine learning in text analysis, in *Handbook of Research on Emerging Trends and Applications of Machine Learning* (IGI Global, 2020), pp. 383–402
19. A. Khan, B. Baharudin, L.H. Lee, A review of machine learning algorithms for text-documents classification. *J. Adv. Inf. Technol.* 1(1), 4–20 (2010)
20. R. Burbidge, B. Buxton, An introduction to support vector machines for data mining. UCL: Comput. Sci. Dept. (2001)

# Assessing Between-Group Differences in Implementation of Lean Six Sigma Management Conception



Slawomir Switek

**Abstract** This work is continuation of research described into two articles: „Empirical evaluation of the revised Technology Acceptance Model for Lean Six Sigma approach – a pilot study” and “Evaluation of the Technology Acceptance Model for Lean Six Sigma approach – the main study” has published by Springer. Due to the fact that acceptance for such key changes as the implementation of Lean Six Sigma is the basic determinant of the long-term success of this type of initiatives, explaining this process requires monitoring it and taking possible corrective actions regarding the change if necessary. Structural Equation Modeling will be used here to assess changes of the acceptance factors over time. This technique is a multivariate statistical analysis one that is used to analyze structural relationships among variables. This paper proposes for the first time a study of the change in the acceptance level over time as a diagnostic tool supporting change management. The prognostic aspect of this study is important for understanding the way changes are made (based on acceptance of six sigma in the organization), because its repetition(s) will help to set the direction of changes. This is the pragmatic value of this work.

**Keywords** Lean six sigma · Total quality management (TQM) · Enterprise resource planning (ERP) system · Technology acceptance model (TAM) · Structural equation modelling (SEM) · Change management

## 1 Introduction

Lean Six Sigma is one of the most popular initiatives in industry sectors, for 10 years it is also being applied in services. Lean Six Sigma approach is a structured and systematic methodology for improving a process focused on reducing process variance as much as possible, minimizing defects (products/services that fall outside of specifications) by using statistics and problem-solving tools intensively [1].

---

S. Switek (✉)  
AWSB University, Dabrowa Gornicza, Poland  
e-mail: [s.switek@gazeta.pl](mailto:s.switek@gazeta.pl)

The number of scientific publications with the terms Six Sigma and Lean Six Sigma has grown moderately in the last 10 years, from about 150 to over 200 publications per year. The number of citations had a much higher dynamics—from approx. 1000 in 2010 to over 3500 in 2020 [2].

Indian researchers note that lean manufacturing and six sigma shown outstanding results. Both the approaches are effective in their own, but with time the organizations implementing these approaches separately may no longer sustain improvements. They are convinced that integrated Lean Six Sigma (LSS) strategy that when implemented will give better results compared to the cases when these two philosophies are implemented individually [3].

The research problem presented in the paper concerns lean management and six sigma methods, which are increasingly used in company management, to increase the efficiency of organization management.

However, there are many challenges in the implementation of Six Sigma [4].

For the successful implementation of the Six Sigma methodology, it is necessary, among others, to have high level of acceptance of these changes by all users participating in their implementation from different levels of management. The examination of the acceptance level of introduced changes is an important tool that allows to evaluate the effectiveness of a comprehensively applied training system and the results of analyzes carried out by improvement teams.

Multiple increases in the value of companies resulting from the application of the methods of business management discussed in the work, concerning international corporations as Toyota Corporation, General Electric or Motorola, have become a common model to be followed by various types of economic organizations and public utility institutions.

The cultural determinants of the success of the application of these management methods made it possible to explain the often-encountered disappointments, especially in the initial stage of implementation, and led to research and analysis of their causes, often leading to a real suspension of the implementation process.

Due to the lack of research on the acceptance of changes based on theories related to perception, which would propose model solutions for the adoption (acceptance) of the Lean Six Sigma conception, an attempt was made to implement them in this work.

Conclusions concerning the factors resulting from perception complete the scope of work on the so-called critical success factors, having their methodical roots in the works of one of the authors of Total Quality Management (TQM)—Ph. Crosby [5].

Research hypotheses often assume that phenomena are shaped depending on certain features of the studied objects. In this way, researchers inquire whether the level of intensity of the studied phenomena or features in the selected groups is the same or different. This study will examine how the acceptance of Lean Six Sigma changes over time in selected subgroups. As a result, managers will be clearly informed about the direction of the change in their organization, and on this basis, they will be able to identify possible remedial measures.

To make the above check, the opinions expressed by the participants and management of the realized implementation processes will constitute the basis for assessing

the effectiveness of the implementation of the Lean Six Sigma conception in this scientific article.

Structural modeling (SEM—Structural Equation Modeling) will be used to perform this task, which allows to estimate the model in several subgroups.

It's basically difficult to find a simple and relevant definition of SEM. Kaplan proposes, that structural equation modeling can perhaps best be defined as a class of methodologies that seeks to represent hypotheses about the means, variances and covariances of observed data in terms of a smaller number of structural parameters defined by a hypothesized underlying model [6].

In the subgroup analysis, it is assumed that the model form is the same in all subgroups and the differences are only in the parameter values. Therefore, the same parameters appear in each subgroup, but their estimated values may be different in each group [7].

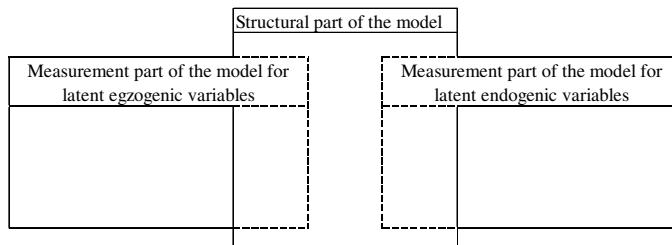
In order to examine the time-changing acceptance of the Lean Six Sigma conception, the customized to the application Technology Acceptance Model, proposed the first time by F.D. Davis in 1985, will be used [8]. Lee et al. define the TAM model as the most commonly used theory describing the individual acceptance of information and technology systems [9].

For this particular study, covariance based (CB), an established variance-based application of SEM, was selected.

## 2 Theoretical Background

In a typical paradigm of correlation studies, referring to the classical use of multiple regression, but extended by the share of latent variables, the following model can be used (see Fig. 1).

The model consists of three parts, and its structural part connects both measurement parts of the model—for latent exogenous variables and latent endogenous ones [10].



**Fig. 1** Basic parts of a structural model. *Source* own elaboration

It is worth to emphasize that the structural part of this model allows testing the basic research hypothesis (based on the  $H_0$  verification, with no mathematical difference), i.e., the hypothesis about the lack of formal grounds for rejecting the proposed theoretical model, if the traditional measure, which is the result obtained in the  $\chi^2$  test, has not exceeded the critical value of the distribution ( $\chi^2$ ;  $p > 0.05$ ). In this situation, the matrix implied by the theoretical model does not differ significantly from the empirical covariance matrix controlled by the system of equations. In other words, the result of the  $H_0$  verification is the basis for accepting or rejecting the research hypothesis: about the admissibility of the causal influence of the psychological reality represented by latent exogenous variables on the reality represented by endogenous latent variables [11].

There are two general ways to investigate group differences with structural equation modeling (SEM). The first method uses a simple path model. When the outcome is a continuous, measured variable, the regression coefficient represents a test of group differences which is equal to a t-test or two-group Analysis of Variance (ANOVA) of the difference between two group means [12].

The second general method of investigating group differences with SEM is to use multigroup models. Multigroup models test separate models in two or more discrete groups [13].

In the study of subgroups, it is assumed that the model representing them has the same form, with the same variables connected by the same relationships. The same set of parameters exists in each group, but their values in each group may be different. Since the form of the model in each subgroup is the same, the path plot for each group will be identical.

The study presented in this article will compare individual parameters between groups. Simple hypotheses, i.e., those relating to the equality of individual parameters, can be verified using the significance test of differences.

The significance test of the difference of two parameters is used to verify whether they are equal in the population. From a technical point of view, any two model parameters can be compared with each other—the same parameter in different subgroups.

The significance of differences test used in this work resembles the t-test for equality of means in independent samples.

The formula for the Student's t-test for independent samples is as follows:

$$T = \frac{X_{bar1} - X_{bar2}}{S_{x1-x2}}$$

$$S_{x1-x2} = \sqrt{\frac{(n1 - 1) * S_1^2 + (n2 - 1) * S_2^2}{n1 + n2 - 2} * \left(\frac{1}{n1} + \frac{1}{n2}\right)}$$

where

T—Student's t-test

$Xbar_1$ —the mean for the first group

$Xbar_2$ —the mean for the second group

$S_1^2$ —variance for the first group

$S_2^2$ —variance for the second group

$n_1$ —the first group size

$n_2$ —the second group size

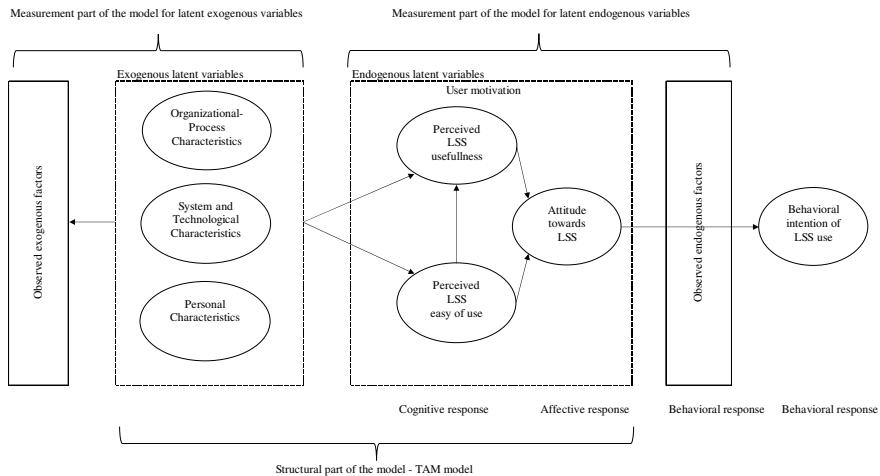
If the value of the test statistic falls into the compartment of  $[-2; 2]$ , then there are no grounds to reject the null hypothesis about the equality of the examined parameters. Therefore, it can be concluded that the tested parameters are equal. If the value of the test statistic exceeds this range, it can be concluded that these parameters (averages) differ significantly [14].

A necessary condition for t-tests is the quantitative (interval or quotient) level of measurement of the dependent variable. In psychology, there is a fairly liberal approach to the definition of the interval scale. It is usually assumed that the quantitative level of measurement has a variable measured using the very popular Likert scale [15].

In the past, in order to make multigroup or between-group analyses, scientists used a variety of approaches, with the partial least squares (PLS) pooled significance test for multigroup becoming the most common. Other researches, especially in areas of social sciences and management, put more emphasis on the use of covariance-based structural equation modeling (CB-SEM) multigroup analysis. According to Qureshi and Compeau [16], both approaches will consistently detect differences under conditions of normality with large sample sizes. With non-normally distributed data, neither technique could consistently detect differences across the groups. In order to deal properly with that kind of issue in the research, the CB-SEM was used together with bootstrapping method and minimum sample size required ( $N > 250$ ) [17].

### 3 Conceptual Model and Research Hypotheses

The research model for the acceptance of the Lean Six Sigma management conception was defined in the previous two works mentioned in the introduction to the article (see Fig. 2). It is a Technology Acceptance Model whose measurement subsystem (set of external factors) has been defined for Lean Six Sigma acceptance. It contains 12 exogenous and 11 endogenous factors.



**Fig. 2** The TAM model adapted to lean six sigma implementation. *Source* own elaboration

As mentioned in Sect. 2, the formal notation of the model will be almost the same as in the case of the model for the entire sample, the only difference will be that individual parameters will be indexed with the symbol of a given group.

The theory of M. Fishbein and I. Ajzen, mentioned later in this work, postulates a causal chain in which behavior is determined by the intentions of showing a certain behavior, intentions result from general assessments or attitudes, and these attitudes are a function of beliefs related to a given individual.

After adopting the attitude, the appropriate behavior of the employee (e.g., using or not using Six Sigma tools or eliminating waste) will be the result. This is of fundamental importance for the entire change, as the implementation of the Lean Six Sigma conception cannot be reduced to the mechanistic application of specific improvement tools or techniques, but should result from the desired behavior. These expected attitudes and behaviors must be absorbed by the organization and included in the company's value system [18].

K. Kmiotek draws attention to the functions of attitudes, namely, she distinguishes their cognitive function—a person adopts a given attitude in order to understand (often simplified) oneself and the world, because all people need reference schemes to better understand reality, and attitudes with a predominant intellectual factor are a kind of guide to facing reality appropriately [19].

The attitudes of the members of the first implementation group, which includes “change agents”, as well as the attitudes of subsequent groups will be crucial for the success of the implementation of the Lean Six Sigma conception.

Due to the above, to determine whether the changes in individual model parameters over time are statistically significant, the following hypotheses will be verified using the t-test described earlier:

Null hypothesis ( $H_0$ ): Means in the compared subgroups are equal;

Alternative hypothesis ( $H_1$ ): Means in the compared subgroups are not equal.

## 4 Methodology of the Main Study

The questionnaire concerning internal and external factors was prepared based on previously made a pilot study.

The whole methodology applied to test acceptance of the Lean Six Sigma conception and specifically, the issues related to the operationalization of variables were previously discussed in the second part—due to the limited publishing volume of this work, we refer here to the dear Reader for further details contained in the scientific article—“*Empirical evaluation of the revised Technology Acceptance Model for Lean Six Sigma approach – the main study*”.

However, it is worth referring to the width of the Likert scale adopted in the study. Here, the 7-point scale was used. F. D. Davis in his first research on TAM model [8] also used 7-point rating scale formats to remain consistent with previous Ajzen & Fishbein recommendations [20].

In terms of scale reliability, the calculated Alpha-Cronbach coefficients for each variable showed values above 0.8 in each case from 0.8270 for personal user characteristics to 0.9247 for attitude towards behavior at the Alpha-level of 0.05. Therefore, it should be assumed that there is no problem with the accepted questions within a given construct. For comparison, in the F. D. Davis’ analysis, the minimum value of 0.96 was achieved.

The chosen scale has been validated and discussed further in Sect. 5 of the work.

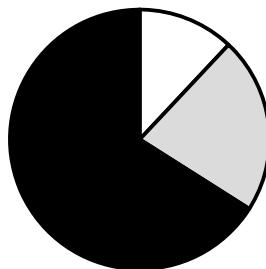
For the chronological order, it should be repeated that the main study was conducted using the CAWI / CATI method in the period of Feb.–July, 2017. The size of this non-random sample was 317 (249 from Poland, 38 from other European countries, 24 from North and South America together with Australia, 5 from Asia, and one respondent from an undefined region).

In order to distinguish between the respondents’ experience in implementing the Lean Six Sigma (LSS) conception, the survey record allows the respondents to be divided into three subgroups (see Graph 1).

- A—people whose experience with the LSS conception was less than 1 year,
- B—people whose experience with the LSS conception was greater than 1 year but less than 3 years,
- C—people whose experience with the LSS conception was greater than 3 years.

The SEM methodology does not make assumptions about how large subgroups should be formed in order to be able to compare them.

**Graph 1** Distribution of respondents by experience in the implementation of the lean six sigma conception.  
*Source* own elaboration



- Group A - less than 1 year
- Group B - more than 1 year, less than 3 years
- Group C - more than 3 years

## 5 Data Analysis and Results

The main research goal was to determine the changes over time of the factors included in the model along with the progressive advancement of the implementation of the Lean Six Sigma methodology.

To compare the subgroups, the program asked to impose test constraints as the model turned out to be unidentified. SPSS Amos used here the method described by McDonald and Krane [21]. Therefore, it was assumed that all regression coefficients for latent exogenous factors COP (organizational and process characteristics), CST (system and technological characteristics), CPU (personal characteristics of the participant of the LSS conception) would be equal for all groups. In addition, all error variances for external factors from O1 to O6, from S1 to S3, from P1 to P3 and the error variances for two endogenous latent variables (PE—perceived ease of use and PU—perceived usefulness) would be equal for all groups.

This assumption makes logical sense, because thanks to this, the endogenous parameters of the TAM model will be calculated under the same external conditions (organizational, process and user attitude).

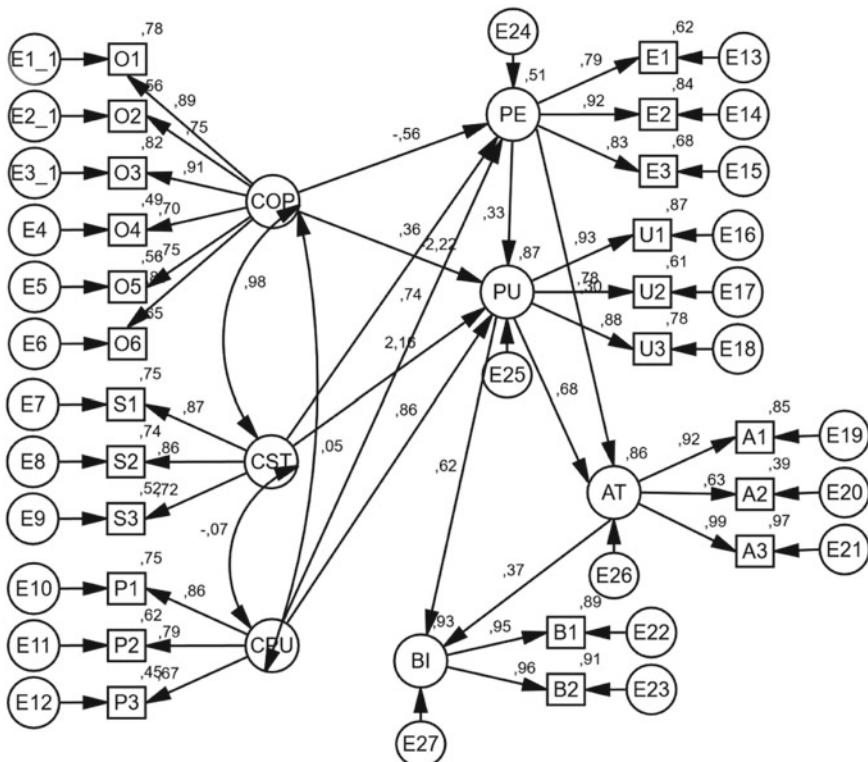
The graphical presentation of the path models for all subgroups is the same as for the entire data set (see Fig. 3).

Using the criterion of time as a determinant of changes in the organization, group A was compared with group B, and then with group C, while group A characterized the respondents least experienced in implementing the LSS conception, and group C the most.

The test statistics of the tests for the significance of differences of analogous parameters in the subgroups showed several values outside the range [-2; 2].

Comparing group A to B it can be observed change for:

- par\_44 and par\_74 (the effect of PU—perceived usefulness on AT—attitude towards behavior); in group A it was statistically significant with a regression coefficient of 0.525, the value of which in group B more than doubled (1.233),



**Fig. 3** Path models of the studied subgroups. Path model for group A—standardized weighting factors. Acronyms used: COP—organizational and process characteristics, CST—system and technological characteristics, CPU—personal user characteristics, PE—perceived ease of use, PU—perceived usefulness, AT—attitude towards behavior, BI—behavioral intention of use *Source* own elaboration

- par\_45 and par\_75 (the effect of PE—perceived ease of use on AT—attitude towards behavior); it was statistically insignificant in both groups and this difference should not be considered further,
- and par\_32 with par\_62 (the effect of AT—attitude towards behavior with A3—using LSS tools is a good idea); this effect in group B decreased compared to group A (change in regression coefficient from 1.1 to 0.864).

Comparing group A to C again a statistically significant difference can be observed for par\_44 and par\_104 (the effect of PU on AT)—regression coefficient in value two times greater in group C (1.034) than in group A, and similar to group B at the same time. The par\_45 and par\_105 (the effect of PE on AT), was still statistically not significant.

The importance of the perceived usefulness (PU) of the LSS conception grew over time, which was to be expected with increasing experience in using these tools and techniques. The user could convince himself of it and confirmed it in this study.

In group B, LSS implementation was rated less as a “positive idea”, which may be related to the already known implementation difficulties at this stage. Group A’s approach could be more idealistic, as a kind of initial enthusiasm was anticipated here.

Moreover, while comparing group A with C, the effect of (par\_29 and par\_89) i.e. PU with U2 (LSS tools are useful in my work) increased significantly, which describes an increase in the value of the regression coefficient from 0.630 to 1.025. This is an interesting observation, as it proves the appreciation of the usefulness of LSS tools at the user’s workstation by the group with the most advanced implementation, i.e., C.

As far as the remaining parameters are concerned, it should be stated that there is no reason to reject the null hypothesis that the influence of these factors is the same in all the examined subgroups.

In the case of the SEM method, the stability of the obtained results should also be assessed. To evaluate the stability of the obtained parameter values, significantly different bootstrap sample sizes and the final form of the path model (with covariances) were used.

If the standardized error is greater than the value of the estimated parameter, then it undoubtedly indicates instability. However, if the estimated parameter converges to a similar value with different starting values, it means that it is very likely that the algorithm has “found” the global minimum of the fit function.

In the analyzed case, the values of the path parameters were generated for the size of the bootstrap sample 250, 1000 and 10,000. Values of estimated parameters and the values of the standardized error S.E. remained the same. For statistically significant parameters, the error value S.E. was a small fraction of the estimated parameter value, which indicates stability. The algorithm of searching for the vector of estimated parameters always ended after 11 iterations, which shows that the global minimum of the fit function was achieved.

At the beginning of the analysis of the obtained data from the respondents’ responses, it was observed that the shape of the histogram is clearly asymmetrical and indicates a left-skewed distribution. Left-hand asymmetry is caused by the responses of respondents who clearly prefer answers lying in the upper part of the adopted Likert scale. There are few answers at the bottom of the scale. Thus, following the respondents’ decision, the scale could be narrowed down by adding the leftmost class to the next one, reducing the asymmetry of the distribution and thus the variability that may affect the stability of the model estimates.

Using the iterative method, two class reductions were made, in which the two left extreme classes were combined, reducing the number of the classes from 7 to 6 and in the next step to 5. This resulted in a change of the scale width from 1 to 6 and from 1 to 5 and the corresponding re-coding the answers. Additionally, after the second-class reduction, the logarithmic transformation of the data was performed (using the natural logarithm) to obtain greater symmetry.

The launched SEM algorithm (with the same sample size and bootstrap sampling) for the estimation of the model parameters did not show a change in the statistical

**Table 1** Comparison of the model parameter values for different Likert scales—main effects only

Relation			For 6-point scale		For 5-point scale		For 5-point scale, logarithmized data	
			Estimate	P	Estimate	P	Estimate	P
PE	←	COP	0.074	0.678	0.085	0.626	0,015	0.929
PE	←	CPU	0.908	***	0.889	***	0.868	***
PE	←	CST	-0.05	0.814	-0.05	0,807	0.024	0.901
PU	←	COP	-0.028	0.834	-0.034	0,795	-0.048	0.711
PU	←	CST	0.018	0.91	0.025	0,875	0.039	0.799
PU	←	CPU	0.625	***	0.631	***	0.736	***
PU	←	PE	0.412	***	0.386	***	0.336	***
AT	←	PU	0.989	***	0.993	***	0.843	***
AT	←	PE	-0.07	0.32	-0.054	0.472	0.003	0.961
BI	←	AT	0.934	***	0.876	***	0.965	***
BI	←	PU	0.209	0.32	0.269	0.161	0.211	0.349

Source own elaboration

significance of the main model relationships, and the values of the path parameters changed to a small extent (see Table 1).

## 6 Discussion and Conclusions

The adopted time criterion made it possible to divide the sample into three subgroups. Model parameters generally did not differ significantly between subgroups, but an increase in perceived usefulness over time was demonstrated, which the respondent could see using Six Sigma tools and techniques. It is also significant that in the group of the most advanced in terms of implementation, the assessment of the usefulness of Six Sigma tools at the user workplace was significantly higher. These are quantifiable indicators that acceptance of the Six Sigma conception is increasing over time. The conducted research and analyzes allow us to conclude that the technology acceptance model is an interesting proposition supplementing the set of diagnostic tools in the field of appropriate change management, which is the implementation of the Lean Six Sigma conception.

The research methodology created for the purpose of this study, which uses hidden factors, assesses their compliance, determines the relationships and examines their change over time, is the methodical value of this work.

The iterations and data transformation made on narrower Likert scales prove that the variability for the entire model is caused by external factors, and the influence of the adopted Likert scale is smaller than it could be expected. This conclusion extends the obtained methodical values of this work.

As long as directions of future research are concerned it is purposeful to repeat the study of changes in the acceptance level in a given organization implementing Lean Six Sigma approach in subsequent years, which would make it possible to monitor this process. The acceptance among users and the consolidation of employees' attitudes are characterized by changes of cultural, internal factors, including interactions in terms of external factors influencing the organization.

Last but not least to sum up, the above studies may not only concern the comparison of simple parameters in subgroups, as presented in this study, but also use the longitudinal methodology that was successfully used in the work [22]. Longitudinal studies allow for an analysis of the variability of a certain phenomenon. They assume the measurement of a variable of interest to the researcher at specific points in time with the same tool and in the same group of people. The changes identified in this way are a function of the differences in the measurement time.

In this research scheme, it is also possible to investigate changes resulting from the interaction between the moments of measurement of specific factors [23].

## References

1. H. Kurnia, H.H. Purba, A systematic literature review of lean six sigma in various industries. *J. Eng. Manage. Ind. Syst.* **9**(2), 19–30 (2021)
2. Z. He, H. Hu, H. Liu, W. Wang, A review on the theoretical research and application of lean six sigma. *Ind. Eng. J.* **24**(5), 1 (2021)
3. S.K. Tiwari, R.K. Singh, S.C. Srivastava, An integrated lean six sigma model for enhancing the competitive advantage of industries. *Recent Advances in Industrial Production*. (Springer, Singapore, 2022), pp. 437–448
4. M.M. Saxena, Six sigma methodologies and its application in manufacturing firms. *Int. J. Eng. Manage. Res.* **11**(4), 79–85 (2021)
5. C. Gastelum-Acosta, J. Limon-Romero, D. Tlapa, Y. Baez-Lopez, G. Tortorella, M.I. Rodriguez Borbon, C.X. Navarro-Cota, Assessing the adoption of critical success factors for lean six sigma implementation. *J. Manufact. Technol. Manage.* **IF7.547** (2021)
6. D. Kaplan, *Structural equation modeling. Foundations and Extensions*. Thousand Oaks, CA: SAGE, 1 (2000)
7. S. Bedyńska, M. Książek, *Statystyczny drogowskaz 3—praktyczny przewodnik wykorzystania modeli regresji oraz równań strukturalnych*. Wydawnictwo Naukowe PWN, Warszawa (2012), pp. 202–215
8. F.D. Davis, A technology acceptance model for empirically testing new end user information systems: theory and results. *Massachusetts Inst. Technol.* 78–79 (1985)
9. G.K. Lee, R.E. Cole, From a firm-based to a community-based model of knowledge creation: the case of the Linux kernel development. *Organ. Sci.* **14**(6), 633–649 (2003)
10. K.A. Bollen, *Structural equations with latent variables*. (Wiley, New York, 1989), pp. 10–39
11. A. Januszewski, Modele równań strukturalnych w metodologii badań psychologicznych. Problematyka przyczynowości w modelach strukturalnych i dopuszczalność modeli. *Studia z Psychologii w KUL*, tom **17**, 218–219 (2006)
12. D. Dimitrov, Comparing groups on latent variables: a structural equation modeling approach. *IOS Press, Work* **26**(4), 429 (2006)
13. K.G. Jöreskog, Simultaneous factor analysis in several populations. *Psychometrika* **36**(4), 409–426 (1971)

14. S. Bedyńska, A. Brzezicka, *Statystyczny drogowskaz—praktyczny poradnik analizy danych w naukach społecznych na przykładach z psychologii*. Wydawnictwo SWPS Academica, Warszawa (2007), p. 197
15. G.A. Ferguson, Y. Takane, *Analiza statystyczna w psychologii i pedagogice* (Wydawnictwo Naukowe PWN, Warszawa, 2003), pp. 197–215
16. I. Qureshi, D. Compeau, Assessing between-group differences in information systems research: a comparison of covariance-and component-based SEM. *MIS quarterly*, 197–214 (2009)
17. J. Nevitt, G.R. Hancock, Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Struct. Equ. Model.* **8**(3), 353–377 (2001)
18. M. Machaczka, Znaczenie miękkiego zarządzania w implementacji szczupłej organizacji. *Zeszyty Naukowe Uniwersytetu Szczecińskiego*, nr 702, *Ekonomiczne Problemy Usług*, no. 87, 199 (2012)
19. K. Kmietek, *Zachowania organizacyjne*. Difin, Warszawa (2012), p. 44
20. I. Ajzen, M. Fishbein, *Understanding attitudes and predicting social behavior* (Englewood Cliffs, NJ: Prentice-Hall, 1980), p. 30
21. R.P. McDonald, W.R. Krane, A note on local identifiability and degrees of freedom in the asymptotic likelihood ratio test. *Br. J. Math. Stat. Psychol.* **30**, 198–203 (1977)
22. L. Zająć-Lamparska, Ł Warchał, M. Deja, Analiza danych podłużnych Modelowanie latentnych krzywych rozwojowych. *Polskie Forum Psychologiczne* **23**(2), 396 (2018)
23. A.D. Farrell, Structural equation modeling with longitudinal data: strategies for examining group differences and reciprocal relationships. *J. Consult. Clin. Psychol.* **62**(3), 477–487 (1994)

# Boosting Blockchain Mechanism Using Cryptographic Algorithm in WSN



P. Anitha Rajakumari and Pritee Parwekar

**Abstract** The bonding and swapping of data within the smart artifices are done by the internet of things (IoT) nowadays. It is assumed that many dangers and attacks on WSN infrastructures are done by the increasing and fast-growing IoT's. As far as the WSN systems are concerned, their most important necessities are observation, reliability, and liability. If such situations have not prevailed, the so-called WSN system could be used by malevolent and hostile users. In the case of the centralized WSN system, when data is conveyed between devices to be shared, problems and risks are quite common. The recent availability of the blockchain and cryptographic tools supports shielding data integrity. This work is mainly framed for studying the temperature and the humidity and various changes seen in the warehousing locations provided for the storage of crops. The data is collected by means of IoT sensor having a wireless sensor network. The blockchain mechanism is employed for the data got from the sensor. We have analysed existing algorithms and its performance. Among the algorithms taught, we have considered the AES to be better performance. So, we have transformed the AES to be a Tuned AES and we have employed it to study the performance in better manner.

**Keywords** Internet of things · Wireless sensor networks · Blockchain · Encryption algorithms

## 1 Introduction

Two main mechanisms are remarkably elaborated in the growth of the Internet of Things. They are sensor networks and wireless sensor networks (WSN) [1]. New problems often arise in the case of safety and confidentiality management for sensor

---

P. A. Rajakumari (✉) · P. Parwekar (✉)

SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India

e-mail: [anitharp@srmist.edu.in](mailto:anitharp@srmist.edu.in)

P. Parwekar

e-mail: [priteep@srmist.edu.in](mailto:priteep@srmist.edu.in)

systems due to specific limitations like low possessions on computation, mechanism of consumption, and hardware of the energy in Wireless Sensor Networks. The research work is divided into two groups; safety and reliability and privacy for sending of data through the network on one side and node documentation and confidence management on another side. These two sets have been discovered for the final ten decades and scholars have brought some solutions. Among these three, security, privacy, and trust management none offers a complete model for content access. But this work particularly focuses on addressing authentication and trust management issues [2, 3].

In today's modern world, most of the research topics are done in the field of microelectronics, network, communication, database centering on wireless sensor networks (WSN) due to its variety of applications. The Wireless Sensor Network combines technology for a variety of purposes such as computing, sensing, and wireless communication. The purpose of using a wireless sensor network is to gather, impart and process the perceived data through its application scenario and the categorization of the hardware is different. By using the WSN data users could collect very interesting information.

The problem mainly arises in the case of WSN is the storage of nodes and it should be rectified. The wireless sensor network relies mainly on a piece of information. The user's concerned about the perception of the data rather than the sensor node and the networks they rely on. The WSN functions efficiently and supplies reliable data storage and is accessible even in a diversified or unreliable environment. Since the repository is constrained, many arduous facets arise in the analysis of exchanging input efficaciously in the cramped space provided for storage.

In WSN, the solutions are mainly based on two mechanisms i.e. mechanism of game theory and mechanism of a reputation for the self-interest problem of nodes. But it is the data processor and the packet forwarding that most researchers focus on. There is no availability of such a specific inducement mechanism of information storage for nodes in Wireless Sensor Networks [4].

The specific determination of using inducement mechanisms based on blockchain is inspiring the network nodes to hoard information and to figure out the problem. The data storage with blockchain technology performs two functions it offers the reliable information storing function and the reward of digit currency in a network node.

The transmitted data from a communication protocol is received and it is secured by the core node of Wireless Sensor Networks and Depending on the information collected it will or will not lead to a reply in [5]. This generic design of WSN is designed to function on general objectives optimistically. One thing which must be improved in the case of their WSN is it has to be improved for its faintness in the security field. Supposing if the data are blocked and alter the transferred information by an attacker or in case it may be missed in some other applications, the sensor nodes' distance may exceed the range of 500 m. In traditional WSN, it is a weak point where centralization meets a lack of propensity and that is the cause for the attackers to rummage the system [6].

## 2 Working Principles of BC

The purpose of using BC in several networks is used to accept the route efficaciously with a similar energy level by using blockchain technology. There are reasons to avoid congestion i.e.

- To increase the productivity of the system
- To raise the well-organized throughput of the system.
- To improve the liveliness of the network.
- To preserve the quality of the services.

### Blockchain Authentication Module (BAM)

Authentication is resolved in networks by a major component like public key infrastructure. A Pretty Good Privacy program and it was synthesized by the Internet Engineering Task Force under the patulous PGP encryption.

\* Encryption with privacy \*Digital Signature verification

\*Uniqueness validation from peers through a web of trust

A new way of attaining the three goals is initiated by the BAM which employs the blockchain as the information processing of backlog public keys cypermated indication and associate intelligence, statistics related to every swelling in the chain is substantiated by approving each element of the network (Table 1).

### Block validity check algorithm

A block is considered to be analytic if it terminates and comprehends a valid (MA)Miner Approval draft provoked by the Miner, demonstrated by the design in the algorithm above. In [7], it is clear that PoW is considered as the consensus algorithm and if it is executed in blockchain it has several disadvantages. Supposing if the hash is to be calculated, a large amount of energy should be consumed. A cost-effective and more well-organized blockchain network has resulted in the case of developing an enhanced consensus algorithm.

**Table 1** Algorithm for BAM

<pre> Require: existing block prior block Ensure: network block legitimacy 1: if (Hash Current Block resolves problem) then Return 1 else 2: return 0 3: end if 4: if MA payload is valid then Return 1 else 5: return 0 6: end if </pre>	<pre> 7: if not Current Block has an event payload for miner NN then Return 1 else 8: return 0 9: end if 10: if not all payloads in the block are valid then 11: return 0 12: end if 13: return 1 </pre>
---	--

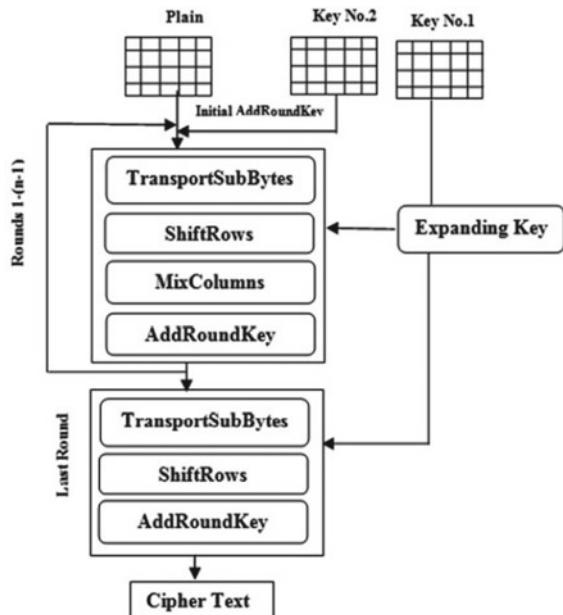
### 3 Proposed Algorithm Tuned AES

The cyber-world is marked by the beginning of ciphering algorithms, and it is initiated by the transformation of information through the internet. There is an establishment of algorithms that find its usage in both the encoding and decoding data for the preservation, drum press, or communication. Advanced Encryption Standard (AES) implies a projecting role among synchronous encryption algorithms. It is also called as Rijndael Algorithm. Because of the high performance of the AES algorithm, it has been nominated as a standard among symmetric cipher algorithms [8]. In this work, we would like to offer a symmetric encryption algorithm. So it is better to prefer a symmetric encryption algorithm. The modification regarding addition or second key is associated with AES. The Modification at the Sub Bytes step is carried out by the addition of transporting Sub Bytes operation. The modified proposed algorithm's performance is done. The overall encryption and description process of our proposed Tuned AES algorithm is shown in Figs. 1 and 2.

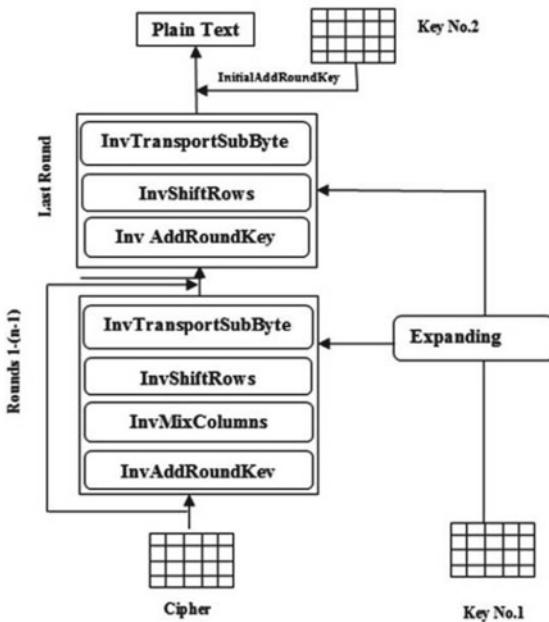
In this work, the following three modifications to the original AES algorithm are done.

**Addition of the Second key:** In the above figures, the operations achieved are as shown and will be added new additional key (Key No: 2). It is assumed that the second key or additional key length can reach upto 2048 bits.

**Fig. 1** Encryption process of tuned AES



**Fig. 2** Decryption process of tuned AES



**Xoring the Second key with plain text:** Before we achieve the key increase encryption process steps, first XOR is an additional key with plain text. This XOR operation is entitled initial AddRoundKey. The output that leads to from the InitialAddRoundKey operation is used as plain text for the following steps. Then the traditional key (key No.1) is expended to generate the sub-keys.

**To modify the SubBytes function:** A new operation called Transport in the original Sub Bytes procedure is included instead of original subbytes operation. And it is retitled as TransportSubBytes. The data are substituted with s-box values after they are transported in the TransportSubBytes operation. In order to get a new state value each element in the state that array (i.e., 8 bits value) is divided into two half 4 bits each and these two halves are get substituted during the transport process. The PHP code employed in TransportSubBytes is shown below.

```

Input: Bytes Arrays
Output: Byte array
Begin
For row 0 to less than 4
For column0 to less than number of bytes
A[row] [column] = Transport of s of row and column
For row 0 to less than 4
For column 0 to less than number of bytes
S[row] [column] = sbox[s[row] [column] & 0 × 000000ff & o&ff Returns
  
```

This process also takes place in the decryption procedure of the Tuned Advanced Encryption Scheme, but the encryption process is inverse. (Inverse Transport SubBytes, Inverse Shift Rows, and Inverse Mix Columns). Finally, the InitialAddRoundKey operation is achieved by XORing supplementary key and the result of earlier processes to recover back the original text.

## 4 Performance Result

The availability of the dataset is seen on the Kaggle website. The encryption algorithm implemented on the Sugarcane crop dataset of Karnataka state, India, and its results is shown in this section. The dataset applied on the crop field is the algorithms AES, Blowfish, tripleDES, DES, and Tunned AES are applied on the crop yield dataset (Table 2).

The following figure shows a comparison graph for encryption and decryption time, energy, throughput for AES, Blowfish, triple des, DES, and Tunned AES. As shown in the graph the throughput for the proposed tunned AES is more and it is showing less energy consumption and time complexity (Fig. 3).

As shown in the above graph the encryption time for the proposed tunned AES is less and also it is showing more encryption time for DES (Fig. 4).

As shown in the above graph the decryption time for proposed tunned AES is less and it is showing more encryption time for Blowfish (Fig. 5).

As shown in the above graph the total encryption and decryption time for the proposed tunned AES is less and it is showing more encryption time for Blowfish (Fig. 6).

As shown in the above graph the throughput for the proposed tunned AES is high and it is showing low throughput for DES and blowfish (Fig. 7).

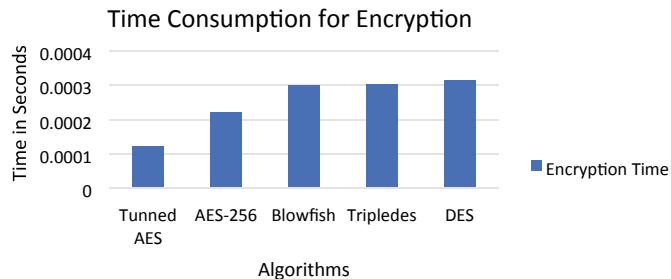
As shown in the above graph the energy consumption for the proposed tunned AES is less and it is showing more energy for DES.

## 5 Conclusion

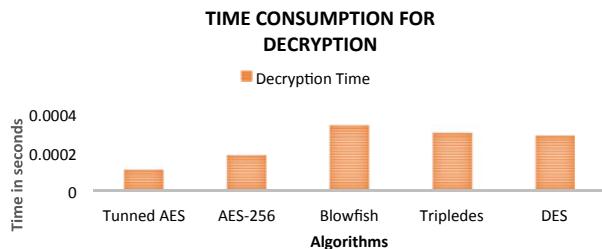
In modern days, the public digital ledger is employed in an unfamiliar approach as a protected, decentralized bin for cryptanalytic keys as well as principle of data in a stately medium. In this paper, it is proposed that the Tuned Advanced Encryption scheme for enhancing security. Many algorithms have been proposed earlier, but the proposed algorithm has given the data security efficiently. The proposed system presents an implementation of Tuned Advanced Encryption Scheme in Blockchain technology over network security. The tuned AES algorithm generates result that are more accurate. The test results show that this proposed algorithm provides good

**Table 2** Comparison table for different cryptographic parameters

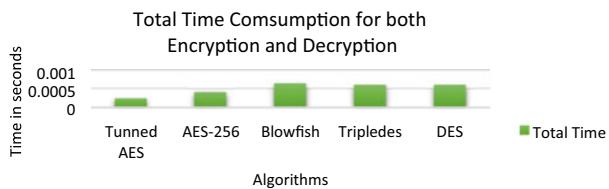
Algorithm	Data size	Key size	Encryption time	Decryption time	Total time	Encrypt on text size	Decrypt on text size	Throughput	Energy
Tunned AES	186	20	0.000121	0.000103	0.000224	280	186	15.3571	0.1513 96
AES-256	186	20	0.000221	0.000179	0.0004	300	186	8.415756	0.2762 68
Blowfish	186	20	0.000302	0.000331	0.000633	268	186	6.162248	0.3772 97
Tripledes	186	20	0.000303	0.000292	0.000595	268	186	6.138006	0.3787 88
DES	186	20	0.000314	0.000278	0.000592	268	186	5.923618	0.3924 97



**Fig. 3** Comparison graph for encryption time

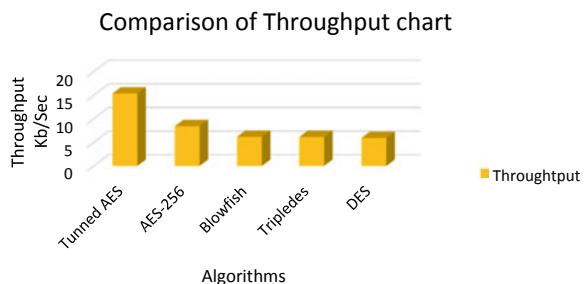


**Fig. 4** Comparison graph for decryption time



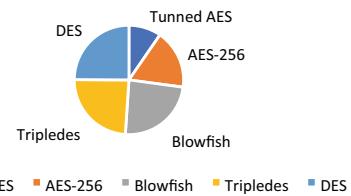
**Fig. 5** Comparison graph for me latency

**Fig. 6** Comparison graph for throughput



**Fig. 7** Comparison graph for energy

Comparison of Energy Consumption Chart



solution to protect the data more secure. Thus, this work concludes with the contributions of experimental study and their performance results on AES, DES, TriDES, BLOWFISH and Tuned AES and all these results have been published.

## References

1. S.R. Anand, R.C. Tanguturi, S. SRD, Blockchain-Based packet delivery mechanism for WSN. Int. J. Recent Technol. Eng. (IJRTE) **8**(2) (2019). ISSN: 2277-3878
2. A. Moinet, B. Darties, J.-L. Baril, Blockchain-based trust & authentication for decentralized sensor networks. arXiv:1706.01730v1 [cs.CR] (2017)
3. A.K.M. Bahalul Haque, M. Rahman, Blockchain technology: methodology, application and security issues. IJCSNS Int. J. Comput. Sci. Netw. Secur. **20**(2) (2020)
4. A.G. Anagnostakis, N. Giannakeas, M.G. Tsipouras, E. Glavas, A.T. Tzallas, IoT micro-blockchain fundamentals. Sensors **21**(8), 2784 (2021). <https://doi.org/10.3390/s21082784>
5. J. Yang, S. He, Y. Xu, L. Chen, J. Ren, A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks. Sensors **19**, 970 (2019). <https://doi.org/10.3390/s19040970>
6. W. She, Q. Liu, Z. Tian, J.-S. Chen, B. Wang, W. Liu, Blockchain Trust model for malicious node detection in wireless sensor networks (2019)
7. H. Sheikh, R.M. Azmathullah, F. Rizwan, Proof-of-Work vs proof-of-stake: a comparative analysis and an approach to blockchain consensus mechanism department of computer science, Prince Sattam Bin Abdulaziz University. Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) **6**(XII) (2018). ISSN: 2321-9653, IC Value: 45.98; SJ Impact Factor: 6.887
8. A.E. Guerrero-Sanchez, E.A. Rivas-Araiza, J.L. Gonzalez-Cordoba, M. Toledano-Ayala, A. Takacs, Blockchain mechanism and symmetric encryption in a wireless sensor network, Faculty of Engineering, Autonomous University of Querétaro, Querétaro 76017, Mexico (2020)

# Virtual Try-On Using Augmented Reality



Aryan Khari, Muskan Gupta, Krishna Mohan, and Saru Dhir

**Abstract** Augmented reality (AR) applications put on another aspect to the customer event by overlaying the users face or factors with virtual items. Virtual 3D take a stab that can give an instinctive and reasonable view for web based shopping and has enormous possible business esteem. The existing 3D virtual take a glance at strategies mostly depend on previous 3D human shapes and piece of clothing formats, which prevents their applications in common sense situations. With this paper we analyse the application non-linear presence and saw the type can clarify positive convincing reactions in direction of AR applications. VTO applications can likewise offer social benefit, permitting purchasers to look for their companions assessments by sharing how an outfit looks on the tweaked model. Despite its utilitarian and liberal qualities, utilizing VTO innovation isn't hazard.

**Keywords** Virtual Try-on · Augmented reality · Digitally modelled · Virtual world · 2D · 3D · Online shopping · Image target

## 1 Introduction

Virtual try-on take an approach at frameworks to make it possible for customers to see themselves wearing various articles of clothing without even wearing them. Actually, wearing the garments can be time costing and frustrating at moments. Such a framework can better the shopping experience by helping clients to make buy choices or if nothing else help to limit selections before actual taking a look at them. Different virtual try on frameworks have been proposed lately [1]. Our strategy is executed utilizing the phenomenon of Augmented Reality (AR), in which digitally modelled pieces of clothing are superimposed on real life customer.

Recently, AR has acquired huge consideration in various arising applications technologies. As an intelligent showcase innovation, AR makes it conceivable to expand the comprehension of online shopping and try on system objects by joining genuine

---

A. Khari · M. Gupta · K. Mohan · S. Dhir (✉)  
Amity University, Noida, Uttar Pradesh, India  
e-mail: [sdhir@amity.edu](mailto:sdhir@amity.edu)

and orchestrated data. The pandemic further provoked the interest for protected, contactless shopping. Because of the propensity of hesitance to up close and personal contact because of COVID-19 alongside the improvement of data [1, 2].

## 2 Literature Review

### 2.1 *Frame of Reference of VTO Innovation*

Online shoppers get hold of choice interaction that includes five stages: issue acknowledgment, data search, assessment, choice and post-buy conduct. Online retailers acquainted VTO innovation with empower online shoppers to choose integral garments from the web-based shop, attempt different matches uninhibitedly and see the outfits on the screen, which helps with their assessment of garments. Thus, this innovation assumes a significant part in the assessment stage, which in turn influences clients buy choice. Without a doubt, it is announced that web-based buyer's disposition towards VTO innovation is identified with social aim towards online retailers or their web based retailing sites, while its qualities and use are related with buy expectation in some ways.

The utilitarian worth of VTO innovation has likewise been surveyed by distinguishing its value and Convenience. Communication with VTO gives an agreeable shopping experience to clients, showing its gluttonous worth. This innovation permits online customers to extend or pivot an outfit, to alter a virtual model utilizing their own body data, or even to transfer their very own image face to make the model look such as themselves.

Earlier examinations have explored VTO application according to a utilitarian point of view or a libertine point of view independently, and a mix of utilitarian and gluttonous point of view, of utilitarian and hazard point of view and of libertine and hazard point of view [4–6].

## 3 Virtual Try-On

A virtual try on takes an approach at its one of the striking highlights of augmented reality. With VTO, a client can virtually attempt a trial on an item prior to purchasing (Try-prior-you-purchase) by just utilizing a gadget. Allowing a client to attempt items in advance of purchasing, takes into account better shopping choices and guarantees that a client gets what he or she wishes.

A VTO take a stab that has totally reformed the manner in which individuals shop on the web. It disposes of the relative multitude of vulnerabilities and questions engaged with web based shopping and clears way for shrewd shopping choices [3, 7].

Brands and organizations across the globe are taking on and executing this component in the entirety of their missions chasing after higher foothold in the market.

Utilizing AR, a client enters the virtual try on experience. This makes it easy to understand experience effectively open from a device. The most awesome aspect of it is that it tends to be sent across stages and channels and subsequently permits brands to divert clients from their sites and online media stages the same to encounter virtual take on at, which further excites the client to make a buy [8, 9].

## 4 Augmented Reality

Augmented Reality is an intelligent innovation that is software based that enhances the clients experience by coordinating extra data (in form of digital assets) into the client's real world. AR gives customers more sensible self-engaging item encounters using super-imposing of assets over them there are by and large two sorts of Augmented Reality Virtual Try On.

### 4.1 2D Overlay AR-Based Try-On

Prior work on 2D overlay try-on take an approach that is for the most portion led in PC designs. The 2D overlay AR based VTO overlays a projected 2D picture of items onto a picture of the purchaser and the genuine climate around the client.

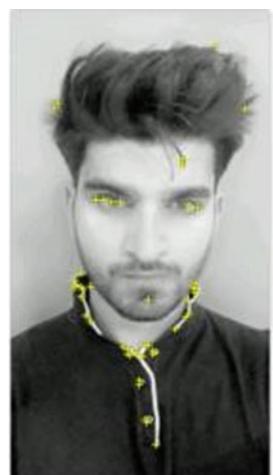
### 4.2 3D AR-Based Try-On

Contrasted with 2D pictures, 3D article of clothing models exactly recreate articles of clothing. Past research zeroed in on coordinating with the 3D virtual article of clothing to client body shape or virtual symbol. As a general rule, the 3D AR-put together attempt with respect which is superior to 2D overlay.

## 5 Proposed Work

### 5.1 Image Target

It addresses pictures that Vuphoria can distinguish and identify (Figs. 1 and 2) The Engine recognizes and tracks the picture by looking at calculated normal elements from the camera picture against a known objective asset information base. When

**Fig. 1** Real photo**Fig. 2** Image with target points

the Image Target is recognized, Vuforia Engine will follow the picture and expand your substance flawlessly utilizing best in market picture identifying innovation. Normal application of Image Targets incorporate collecting and expanding printed media and item bundling for advertising efforts, gaming, and picturing items in the climate where the item was planned to be utilized [10, 11]. The Image Targets test application addresses Vuforia's centre elements and can be an incredible beginning stage to assist you with carrying out and arranging Image Targets in an application.

## 5.2 *Visualization*

The Visualization unit is the last and the most essential segment of the whole cycle.

The product used to accomplish this assignment: Unity Engine and Vuforia motor. The Unity motor is utilized to make augmented reality and expanded reality reenactments. (Fig. 3) An Android or IOS Based device, Augmented Reality (AR) perception application is planned and delivered in Unity, shows the representation plan for the transport engine well being [12].

There can be a few machines in a plant and, the AR perception of sensor information will be diverse for all. The application should associate with a specific part of the cloud to see the boundaries of a particular machine. Vuforia motor purposes this issue by utilization of Markers and Image Targets. As indicated by, Markers are the pictures that Vuforia utilizes as the perspectives to show our projections. They are novel and connected to just a solitary article. Picture targets are the component focuses removed from the Markers. Here a QR code goes about as the Marker and, is taken care of to Vuforia motors information base. Vuforia first identifies the picture focus, as and contrasts the component focuses and the getting edge of Marker through the gadget camera [13].

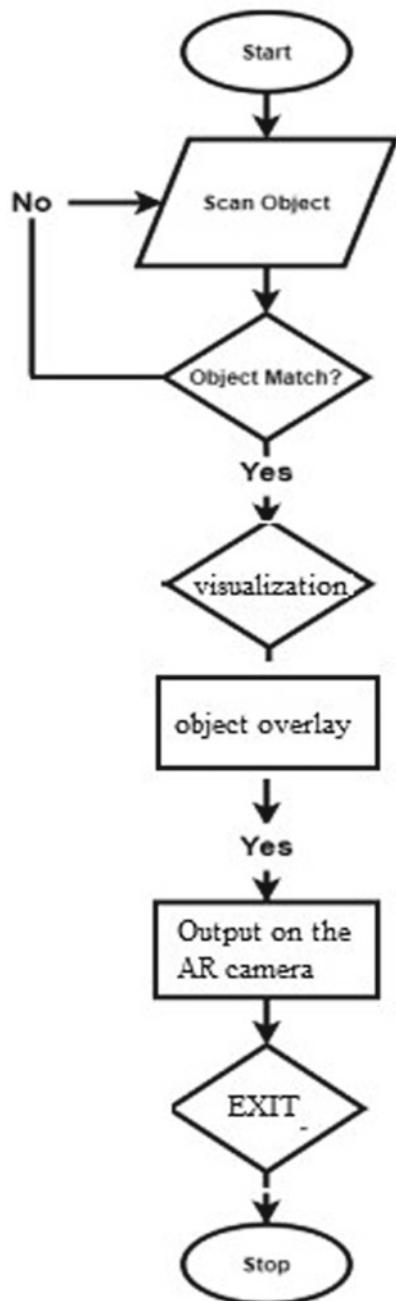
# 6 Software and Vuphoria Engine

## 6.1 *Unity*

Unity has assets to rejuvenate your vivid vision. Industry-driving programming advancement stage sets devices reason worked for Augmented Reality makers and a brought together work process across gadgets that allow us to zero in on stretching the boundaries of your creative mind.

A structure reason worked for AR improvement permits you to create your application once, then, at that point, convey it across different versatile and wearable AR gadgets. It incorporates centre highlights from every stage, just as remarkable Unity includes that incorporate photograph practical delivering, physical science, gadget improvements, and then some. (Fig. 4) AR Foundation allows you to take at present inaccessible elements with you when you switch between AR stages. In the event that an element is empowered on one stage however not another, we put snares in so it's all set later.

**Fig. 3** Process cycle of VTO





**Fig. 4** Image target in unity work space

## 6.2 *Vuphoria*

The Unity Editor is a famous and valuable writing stage to make state of increased reality encounters for both handheld gadgets and advanced eyewear. After establishment of Unity, the Vuforia Engine is effectively added to any project, and our Core Samples are accessible in the Unity Asset Store to kick you off. Vuforia Engine is a product improvement pack for making AR based applications. Engineers can without a doubt add advanced PC vision convenience to any application, allowing it to see pictures and dissent, and speak with spaces in all actuality.

## 7 Algorithmn and Analysis

### 7.1 *Model Generation*

Previous examination on piece of clothing demonstrating began from the 2D plan design or 2D portrayals. While different techniques investigated piece of clothing resizing and move from 3D format pieces of clothing. Contrasted with 2D plan designs, 3D attire of clothing models mimic the article of clothing all the more exactly. Subsequently, our technique stretches out these strategies to plan the 2D picture to the 3D attire of clothing models. The surface planning technique builds the authenticity of the article of clothing.

<b>Fig. 5</b> Augmentable rating	Type: Single Image
	Status: Active
	Target ID: 4d5d82de0cf0423eb4ba0dc5e7d787bc
	Augmentable: ★★★★☆

## 7.2 Vuforia Database

The motivation behind a Device Database additionally mentioned to as datasets is to give the AR application a locally available information base of targets that is set by key points. Various data sets to incorporate with your Vuforia Engine application are present. The Vuforia Target Manager permits you to make and oversee Device Database and the objectives inside them before they are downloaded.

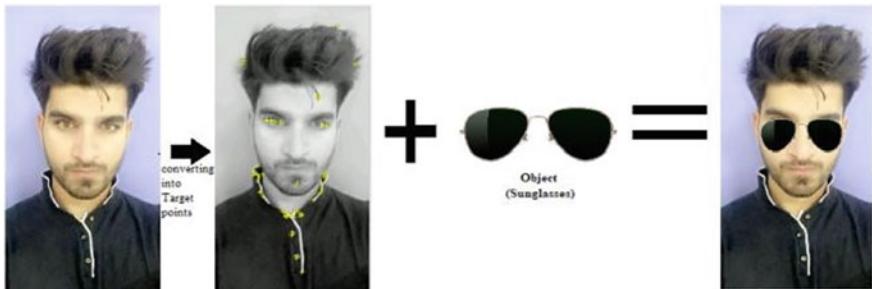
The Target Manager permits you to make and deal with those information bases on the web, and gives download choices to both Unity and local activities. Item Targets, Image Targets, Multi Targets and Cylinder Targets are made and overseen utilizing the Target Manager. The suggested most extreme size for a Device Database is 1000 picture based targets, however it is feasible to help a bigger number of picture put together targets depending with respect to the pictures utilized.

## 7.3 Augmentable Rating

An AR rating characterizes how amazingly a picture is distinguished and followed utilizing the “Vuforia”. This rating is shown in the Target Manager and is returned for each transferred target through the web API when utilizing Databases. The AR can go from 0 to 5 for some random picture. The star rating results the power of the object. The star rating can change somewhere in the range of 0 and 5 stars for 0 being a hard to identify image and 5 being an easily targetable image (Fig. 6). A higher rating is better considering the way that the picture target gives more straightforward identification and more steady following (Fig. 5).

## 7.4 Point Based Analysis

The Engine distinguishes and tracks the picture by looking at marked normal highlights from the camera picture against a known objective asset information base. When the Image Target is distinguished.



**Fig. 6** Result of VTO

## 8 Results and Conclusion

The Augmented reality based Virtual Try on was propitiously developed to make it possible for customers to see themselves wearing various items without even wearing them. AR makes it conceivable to expand the comprehension of online shopping and try on system objects by joining genuine and orchestrated data. The pandemic further provoked the interest for protected, contactless shopping. Because of the propensity of hesitation to up close and personal contact because of COVID-19 alongside the improvement of data. As a matter of fact wearing the garments can be time costing and frustrating at moments. Such a framework can improve the shopping experience by helping clients to make buy choices or if nothing else help to limit selections before actual taking a look at them. The real life image was marked with targets using Vuforia engine. The article was successfully superimposed on the picture target.

## References

1. Y.A. Sekhavat, Privacy preserving cloth try-on using mobile augmented reality, *Multimedia IEEE Trans.* **19**(5), 1041–1049 (2017)
2. J. Zhi, R.Xie, H. liu, X. Wang, G. Yang, P. Peng, X. Luo, Y. Chen, A survey of interactive technology in augmented reality, in *Information Technology and Computer Application (ITCA) 2020 2nd International Conference on* (2020), pp. 101–104
3. A.R. Smink, E.A. van Reijmersdal, G. van Noort, P.C. Neijens, Shopping in augmented reality: the effects of spatial presence, personalization and intrusiveness on app and brand responses. *J. Bus. Res.* **118**, 474–485 (2020)
4. A. Mir, T. Alldieck, G. Pons-Moll, Learning to transfer texture from clothing images to 3d human (2020)
5. X. Liang, K. Gong, X. Shen, L. Lin, Look into person: joint body parsing & pose estimation network and a new benchmark. *IEEE Trans. Pattern Anal. Mach. Intell.* (2018)
6. M.R. Minar, T.T. Tuan, H. Ahn, P. Rosin, Y.-K. Lai, Cpyton+: clothing shape and texture preserving image-based virtual try-on, in *CVPRW* (2020)
7. P.B. Dweep, S. Dhir, M. Hooda, Impact of IOT in current Era, in *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (IEEE, 2019), pp. 334–339. ISBN: 9781-7281-0211-5

8. B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng, NeRF: Representing scenes as neural radiance fields for view synthesis, in *Proceedings of European Conference on Computer Vision (ECCV)* (2020)
9. M. Atzmon, Y. Lipman, SAL: sign agnostic learning of shapes from raw data, in *Proceedings of Computer Vision and Pattern Recognition* (2020)
10. Z. Huang, Y. Xu, C. Lassner, H. Li, T. Tung, ARCH: animatable reconstruction of clothed humans (2020)
11. S. Sridhar, S. Sanagavarapu, Instant tracking-based approach for interior décor planning with Markerless AR (2020)
12. S. Dhir, D. Kumar, Automation software testing on web based application. *Adv. Intell. Syst. Comput.* **731** (2018). Springer, ISBN: 978-981-10-8848-3
13. Y. Desai, N. Shah, V. Shah, P. Bhavathankar, K. Katchi, Markerless augmented reality based application for e-commerce to visualise 3D content, in *Inventive Research in Computing Applications (ICIRCA) 2021 Third International Conference on* (2021), pp. 756–760

# Fractures Classification and Detection Using DIP and ML



Nishit Dhamija and Saru Dhir

**Abstract** In modern years, the identification and categorization of fractured bone has been an extensively debated field and various scholars have anticipated various models to solve this setback. A good collective methodology to sort all types of fractures that occur in body hasn't been outlined yet. The article targets to investigate and assess the kinds of fractures that occur in humans through DL and DIP. This paper uses a different approach to recognize bone fractures. For easy diagnosis of types, we are using only X-ray, technology available in hospitals and classifying them through machine learning.

**Keywords** X-Rays · CNN · Deep learning · Machine learning

## 1 Introduction

Bone injuries and fractures are very common and happen to thousands and millions of people every day. They are most commonly caused by either car accidents, sports injuries, falls, etc. and take quite a time to heal properly.

There are various approaches those are cast off to distinguish the fracture such of them are as follows [1, 2]:

- **X-rays:** It is a very high radiating electromagnetic ray which produces two-dimensional picture of the damage happened.
- **Bone scan:** A bone scan is generally used to find fractures that don't show up on an X-ray.
- **CT scan:** It uses computer and X-ray to craft a detailed cross-section of bone.
- **MRI:** It creates a very detailed image using strong magnetic field, gradient and radio waves.

---

N. Dhamija (✉) · S. Dhir  
Amity University, Noida, Uttar Pradesh, India  
e-mail: [nd6938@gmail.com](mailto:nd6938@gmail.com)

Apart from the breakage bone fractures sometimes also becomes a reason for the following complications in the body such as blood clots, cast-wearing complications, compartment syndrome, hemarthrosis.

The symptoms of fracture may vary depending upon its location, patient's age, health, and injury severeness [3]. However, people with bone fractures usually have some of the following symptoms:

- ache
- swollen area
- bruising
- not able to put or sustain any weight on the injured area In severe cases, a patient may experience:
- wooziness
- indistinctness
- repugnance.

ML is the technique of AI which involves procedures to parse data, learn from it, then put together well-versed decisions to apply what they have learned [4]. AI algorithms programmed for continuous learning that simulates that of a virtual personal assistant.

## 2 Literature Review

With few exceptions, surgeons have neglected to perfect their records upon the essential points like, exact description of initial lesion, period of hospital residence, injury and recovery time gap, degree of permanent deformity and loss of function. All surgeons must admit the necessity of collecting data for future guidance in giving the best prognosis and treatment. But it is probable to make this material of value only by carefully classifying and correctly collecting large numbers of cases, so that we can find the normal average duration of disability which is our best guide for treatment in a given type of case.

So, here we review the studies devoted to detecting fracture cases (Table 1).

## 3 Methodology

In this segment, we have described the whole system plan which incorporates, image preprocessing, image dissection and fracture-detection [6]. Figure 1 demonstrates the flow schema of a digital image processing. Preprocessing image segmentation gives the smooth, noiseless and sharpen image Vertex AI available on the GCP was primarily cast-off to sequence the model. Vertex AI embroils administered learning chores to triumph an anticipated result. The particulars of procedure and drilling methods alter grounded on data type (Fig. 2).

**Table 1** Literature review table [1–3, 5–12]

S. No	Paper Title	Author and Year	Method Used	Accuracy	Comments/observations
1	Analysis on leg bone fracture classification and detection using X-ray images	Myint et al. [1]	HCD, DT, KNN	85%	In this, HCD is used to locate cracked points DT is cast-off to sort image as fractured or unfractured KNN is cast-off to identify the type of fracture
2	Bone fracture detection using edge detection technique	Johari et al. [2]	Sobel operator, CED	87.3%	Sobel operator is cast-off to improve the efficiency of system and it diagnoses the hairline fracture more effectively
3	Novel approach for bone fracture detection using image processing	Dhiraj et al. [3]	SVM, KNN	85%	The different classifiers like SVM, KNN are used for image categorization
4	Detecting leg bone fractured X-ray images	Myint et al. [5]	CED	—	Much higher accuracy can be achieved by using high-res images
5	Bone fracture detection using OpenCV	Kurniawan et al. [6]	CED using OpenCV	66.7%	Accuracy of system is effected by image quality. Better the image quality, better the results
6	Dermatological disease detection. using DIP and ML	Vinayshekhar [7]	kth Nearest Neighbor (kNN), Decision Trees (DT) Artificial Neural Networks (ANN)	95%	The system has two Stage refinement process and merging the two stages increases the accuracy and efficient in the dermatological disease detection. Similar can be adopted for identification of bone fractures
7	Machine learning on mobile: an mobile app for skin cancer detection	Dai [8]	Convolutional neural network (CNN) using TensorFlow	75.2%	The system can be modified for fracture recognition in android device using Tensorflow lite

(continued)

**Table 1** (continued)

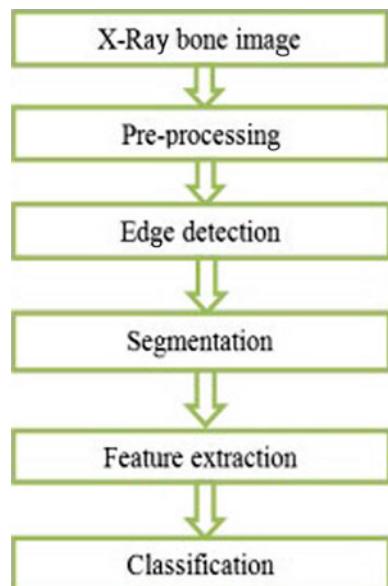
S. No	Paper Title	Author and Year	Method Used	Accuracy	Comments/observations
8	Fracture detection in X-ray images through feature fusion of stacked random forest	Cao et al. [12]	Random forests algorithm (RFA) & Support Vector Machine (SVM)	81.2%	SVM and single layer random forests increases its effectiveness
9	Multiple classification system for fracture recognition in X-ray images	Umadevi et al. [11]	SVM, BPNN, KNN	SVM Accuracy- 91.89 BPNN Accuracy-90.46 KNN Accuracy-89.76	Experimental output displays that cooperative model that combines BPNN, SVM and KNN with both the features significant improvement in relation of exactness
10	Determining type of long bones fractures in X-ray images	Al-Ayyoub et al. [10]	Support Vector Machine (SVM) Decision Tree (DT) Nave Bayes (NB)	85%	SVM was most accurate than other ones Integrate this method in software of an x-ray machine is very efficient in teaching and research purposes
11	Diagnosis of skin diseases using convolutional neural networks	Rathod and Jainesh [9]	Convolutional neural network (CNN)	90%	CNN technique can be cast off for diagnosing the bone fracture and same can be cast off for the classifiers

An image dataset was used train, test, and validate the ML mode1 with pictures which are labeled for classification and object detection.

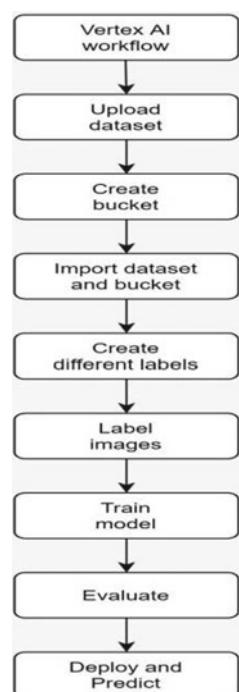
## 4 Dataset

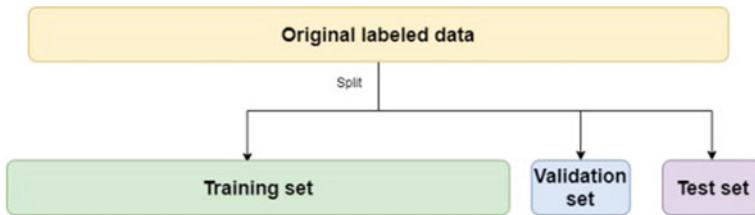
A Dataset of total of 1509 images where 712 X-ray images were that of healthy bone, 388 were of open fracture & 409 X-ray images that depicted close fracture. This is a dataset which was composed from Dr. Gupta in Life Hospital, different orthopedic labs (Max lab, Ganesh diagnostics, Orbit imaging), and some of the pictures are amassed from open source kaggle.com and istockphoto.com. If we don't identify

**Fig. 1** Image processing flowgraph



**Fig. 2** Vertex AI work flow





**Fig. 3** Data split

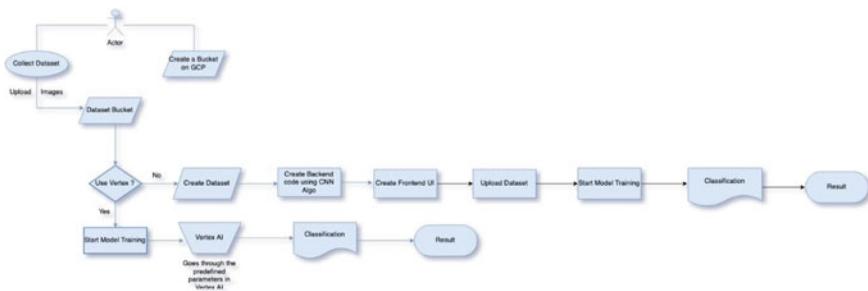
the split, then Vertex AI involuntarily uses 80% of our images for training, 10% for testing, and 10% for validating.

The dataset was splitted into three parts; 1205 pictures were implement for guiding the model, 152 pictures were implement for model validation & 152 pictures were implement for model testing (Fig. 3).

## 5 Workflow

The workflow of our project started with assembly of primary datasets comprehending three image classes: one belonged to X-rays of healthy bone and one of images belonging to the closed fracture and the third belonging to the open fracture. The dataset consist of three parts; 1205 pictures were implemented for instructing the model, 152 pictures were implemented for model validation & 152 pictures were implemented for testing the model. The resulted dataset was cast-off to instruct model in the subsequent phase. After teaching, the model was verified for its functioning in detection of fractures.

Vertex AI implemented to train & validate the model. After model validation was completed the, it was deployed and tested. To test the deployed copy a batch prediction dataset was created, resulting in exactness of 78% with a recall value of 73.1% (Fig. 4).



**Fig. 4** Work Flow

The following steps detect how we start:

1. First create the dataset
2. Then write backend code using CNN algorithm
3. After this create frontend UI
4. Upload dataset for training
5. Start training the model
6. Predict the outcome with model.

## 6 Result

In the proposed model we present a machine learning architecture based on supervised learning for fracture detection using X-ray images dividing them into 3 classes, Open fracture, Closed fracture and Healthy bone based on the dataset. The model was educated using Vertex AI which is a google cloud based unified UI for the entire ML workflow. The computer is fed with images that can help in the precise identification of the bone. The system is fed with hundreds of images of bones. The model thus creates and learns about different features that make up a fractured bone or differentiates it from other images. An inbuilt cross validation was implemented within Vision AI to get the achievability of consuming bone images to diagnose the fracture. The initial results express promising. The validation precision and average precision for Vertex AI is measured as 77.9% and 0.849 respectively.

The following table depicted the precision values using the Vertex AI technology (Table 2).

The graph in Fig. 5 spectacles the tradeoff amidst precision and recall at different confidence levels in Vertex AI.

The graph in Fig. 6 displays how our model performs on the top-scored label along the full range of confidence threshold values in Vertex AI. A higher confidence threshold produces fewer false positives, which increases precision. A lower confidence threshold produces fewer false negatives, which increases recall.

Figure 7 displays how often model categorized each label accurately (blue), and which labels were most frequently jumbled for that label (grey).

**Table 2** Precision values

Average precision	0.849
Precision	77.9%
Recall	73.1%
Total Images	1509

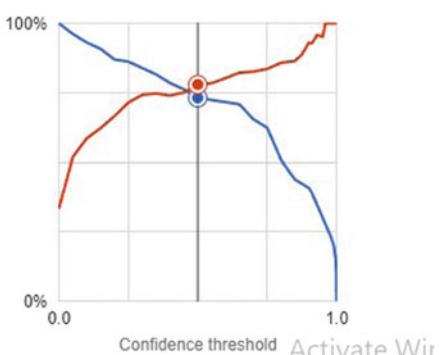
**Fig. 5** Precision recall curve vertex AI

Precision-recall curve



**Fig. 6** Precision recall curve vertex AI

Precision-recall by threshold



**Fig. 7** Confusion matrix using Vertex AI

True label	Predicted label		
	Closed_Fracture	Healthy_Bone	Open_Fracture
Closed_Fracture	59%	38%	3%
Healthy_Bone	4%	89%	7%
Open_Fracture	—	22%	78%

## 7 Conclusion and Future Scope

In this paper, a system is developed to distinguish and sort the fracture type using ML (deep learning) and digital image-processing techniques. Plying X-Ray image as an input image, it can detect and classify the fracture. The X-Ray images were gathered from distinctive sources, and to overcome issue of small dataset in deep learning, a large dataset was assembled which consisted of 1509 images of healthy and fractured bones combined. The overall accuracy came out to be as 77.9% in which the accurateness of the healthy bone resulted as 77.8% and closed fracture as 84.2% and open fracture as 88.4% which was tested using Vertex AI. It can be further improved using the CNN algorithm.

For now, model is verified for the finger, palm, wrist and arm bones, which if we consider for future scope can be further trained and prepared for all other bones present in the body. Also, the bone classification can be further expanded to distinctive forms of closed and open fractures.

## References

1. W.W. Myint,, K.S. Tun, M. Tun, Analysis on leg bone fracture detection and classification using X-ray images. *J. Mach. Learning Res.* **3**(3), pp. 49–59 (2018)
2. P. Srinivasulu, J. Vamsi, K. Drutesh, K. Prudhvi, Bone fracture detection using image processing. In: Electronics and Communication Department, GITAM School of Technology—Bengaluru, Vol. 5, no. 6 (2020).. ISSN: 2455–2631 June IJSDR
3. S. Sapra, M. Hooda, A. Chhabra, S. Dhir, Smart med-minder. *Indian J. Publ. Health Res. Dev.* **9**(11), 661 (2018)
4. S. Jojo, S. Dhir, A. Madhurima, A machine learning APP for prediction of pandemic such as COVID-19. In: Advanced Computing Techniques: Implementation, Informatics and Emerging Technologies, Bentham Publishers, pp.27–38 (2021)
5. S. Myint, S. Khaing, H. Tun, Detecting leg bone fracture in X-Ray images. *J. Sci. Technol. Res.* **5**(06) (2016)
6. F. Kurniawan, K. Darma, K. Sudana, Bone fracture detection using OpenCV. *J. Theoret Appl. Inf. Technol.*
7. K. Vinayshekhar, Dermatological disease detection using DIP and ML (2016)
8. X. Dai, Machine learning on mobile: an mobile app for skin cancer detection (2019)
9. K. Rathod, S. Jainesh, Diagnosis of skin diseases using convolutional neural networks (2018)
10. Al-Ayoub, Determining type of long bones fractures. *Jordan Univ. Sci. Technol.* **4**(3) (2018)
11. N. Umadevi and S. N. Geethalakshmi. Multiple classification system for fracture recognition in x-ray. pp. 1–8, doi: <https://doi.org/10.1109/ICCCNT.2012.6395889> (2012).
12. Y. Cao, H. Wang, M. Moradi, P. Prasanna, Fracture detection in x-ray images through stacked random forests feature fusion. 801–805 (2015). <https://doi.org/10.1109/ISBI.2015.7163993>

# Sentiment Analysis in Social Media Website



Aditya Rao, Mehak Khanna, Priya Sinha, and Saru Dhir

**Abstract** Sentiment analysis is the automatic extraction of positive or negative attitude expressions from textual information, and it has gotten a lot of interest from scholars in the last decade. Additionally, the popularity of online review sites, social networks, and personal blogs has grown rapidly in line with the advancement of technology; the users actively use on-line review sites, social networks, and personal blogs to express their opinions. They express people's, organisations', locations', events', and notions' positive and negative sentiments. Natural language processing and machine learning, as well as other ways to working with massive amounts of text, have made it possible to start extracting sentiments from social media". In this paper we will show the design of our social and media website along with different sentiment analysis algorithms like Naïve Bayes, LSTM and VADER; the analysis of these different algorithms will help us conclude for the most suitable algorithm to be applied in the social media website.

**Keywords** Social media website · Sentiment analysis · Naïve bayes · LSTM · VADER · NodeJS · MongoDB · MVC architecture

## 1 Introduction

When people live in society, they develop opinions about the world around them whether good or bad. They form opinions about people, products, places, and situations that are both favourable and bad. Attitudes like these are referred to as sentiments. The study of approaches for extracting sentiments from written languages is known as sentiment analysis [1]. "The surge of publicly available, user-generated text on the World Wide Web has stemmed from the growth of social media". This data and information have the ability to deliver real-time insights into people's feelings. Websites like Facebook, Instagram, Twitter, and others may capture the opinions or word of mouth of millions of consumers. In machine learning and network analysis,

---

A. Rao · M. Khanna · P. Sinha · S. Dhir (✉)

Department of Information Technology, Amity University Noida, Noida, India  
e-mail: [sarudhir@gmail.com](mailto:sarudhir@gmail.com)

communication and the availability of these real-time perspectives from individuals or citizens all over the world has ushered in a revolution. “People are more willing and fuller of content than ever before to share details about their lives, knowledge, experiences, and ideas with the wider world via social media. They actively participate in events that occur in society by expressing their ideas and making statements. People have used social media to share their area of expertise and emotions, prompting them to not only give positive feedback but also to troll excessively”. Here we analyse Naïve Bayes, LSTM—Linear Regression and VADER. On performing analysis of these algorithm, we managed to get information and a comparison between them. For the social media website, we used MVC architecture and NodeJS for backend. For database we used MongoDB.

As a result, the increasing demand for the features of sentiment analysis on social media websites necessitates the development of a tool that can analyze sentiments on our social media while addressing trolls and bullying, and then develop a website for users where they can post, like or comment about anything they want without being anxious of being trolled or bullied.

## 2 Literature Review

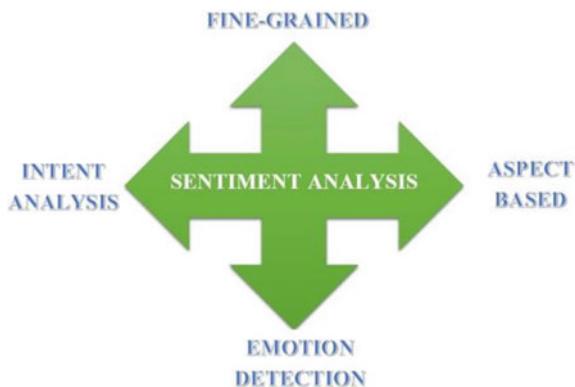
Sentiment analysis is a machine learning technique which is mainly used to classify or to review text or corpus.

Sentiment analysis is basically stated as opinion mining technique [1]. It uses the machine learning, Natural Language Processing (NLP) which recognizes the particular emotion tone is the text belongs to whether it is positive or negative. It includes numerous other techniques specifically-Artificial Intelligence, Data Mining and machine learning to mine corpus or text to identify sentiment. Sentiment Analysis is commonly used by different organization to review their product and services as the sentiment analysis algorithm mitigates the need of manual data processing.

Sentiment analysis further branches into four distinguished types and this segregation is based on different purposes (Fig. 1).

- **Fine-grained Sentiment:** This type of analysis gives report of customer feedback. The accuracy of result is judged by the input in the quadrant of polarity.
- **Aspect Based:** It is used by the organization or companies for their particular product or service so it analyzes customer review or feedback on different parameters or features of the company’s product. It basically securitizes product’s attributes or aspects to ensure accuracy.
- **Intent Analysis:** Intent Analysis is used to extract the deeper knowledge of customer’s intention [3]. It basically categories customer by securitizing their intentions on that particular product.
- **Emotion Detection:** It is considered as the most sophisticated algorithm of sentiment analysis. It uses machine learning and lexicons to study and determine the

**Fig. 1** Types of sentiment analysis [2]



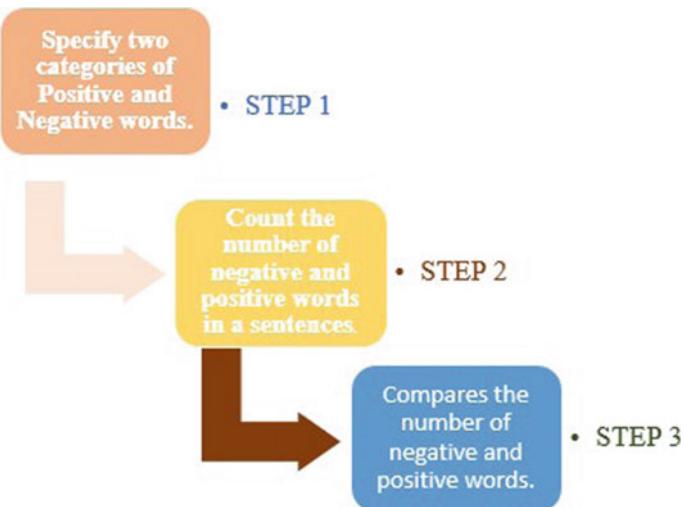
polarity of sentiment. Lexicons are the set of words it can be either be positive or negative words. Sentimental analysis broadly has three approaches.

- **Ruled-based approach:** In this approach there are pre-defined set of rules which is used to determine polarity, subject and topic of an opinion. There is an extensive use of NLP techniques such as Lexicon, Stemming, parsing and tokenization [4]. It is completed in following steps (Fig. 2).
- **Automatic approach:** It is not worked as a manual processing or opinion mining. In this a filter used technique of neural network which works in such a way that the filter the classifier is given a set of corpuses and it returns the category such as Positive, negative and neutral.
- **Hybrid Approach:** Hybrid is basically combination of both ruled based and automatic approach. In this a predefined dataset of corpus which is fed to the classifier filter to train the system, and then with the assistance of neural network techniques such as linear regression or LSTM (Long short-term memory) it classifies the corpus in a particular sentiment.

### 3 Methodology

#### Design

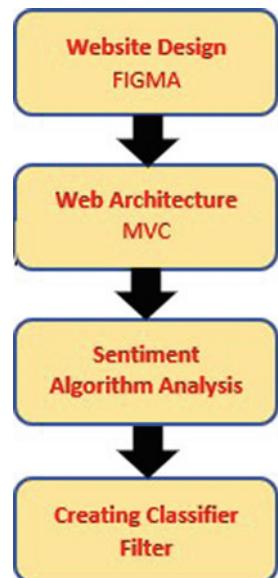
The website is designed with Figma. The user interface (UI) Figma is a cloud-based design application that is similar to Sketch but has some key distinctions that make it superior for team collaboration [5]. Real-time file collaborative effort aids in the reduction of “design drifting”, which is defined as misinterpreting or deviating from an agreed-upon design. Design drifting typically occurs when an idea is conceived and quickly implemented during the course of a project. Unfortunately, this frequently leads to deviations from the established design, resulting in friction and rework (Fig. 3).



If the number of positive word appearances is greater than the number of negative word appearances, the system returns a positive sentiment, and vice versa. If the numbers are even, the system will return a neutral sentiment.

**Fig. 2** Steps of ruled based approach [3]

**Fig. 3** Dataflow diagram of the proposed system



### MVC Architecture

Model-View-Controller (MVC) is an abbreviation for Model-View-Controller. The MVC pattern divides the application and its development into three interconnected parts.

**Model:** A model is a representation of the structure of data, its format, and the curtailment under which it is stored. It stores the application's data. Essentially, it is the application's database.

**View:** The view is what the user sees. Views reap the benefit of the Model to present data in the format desired by the user. A user may also be permitted to modify the data presented to him or her. They are made up of static and dynamic pages that are rendered or sent to the user when they make a request for them.

**Controller:** The controller manages the user's requests and then generates the adequate reaction, which is fed to the viewer [6]. The user "typically interacts with the View", which generates the appropriate request, which is handled by a controller. As a response, the controller renders the appropriate view using the model data (Fig. 4).

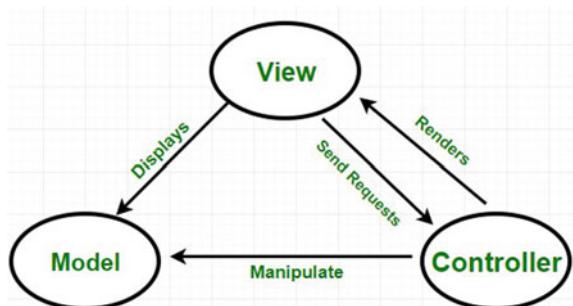
### NodeJS Server Architecture

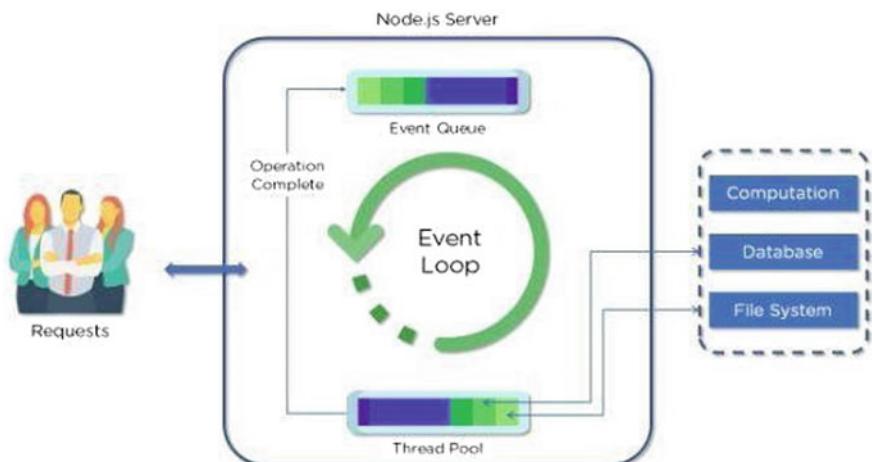
To handle multiple concurrent clients, Node.js employs the "Single Threaded Event Loop" architecture. The Node.js Processing Model is built on the JavaScript event-based model in conjunction with the" JavaScript callback mechanism (Fig. 5).

### Implementation of Text Classifier Filter

For the implementation of the text classifier filter for Sentiment Analysis Word Count Vector method is used in which every word is placed in a column and have a frequency count from the dataset of the model. To plot or to represent the relative value or importance of a word in the whole dataset or corpus method of TF-IDF vectorization is used. It basically analysis the importance of a particular word by computing the occurrence of each word in dataset. After analysis of each word to further increase the accuracy of the filter positions of words is also checked in the vector space by examine the text corpus and all the neighbouring words that surrounds that keyword this method is know as word embeddings.

Fig. 4 MVC architecture [6]





**Fig. 5** NodeJS server architecture [7]

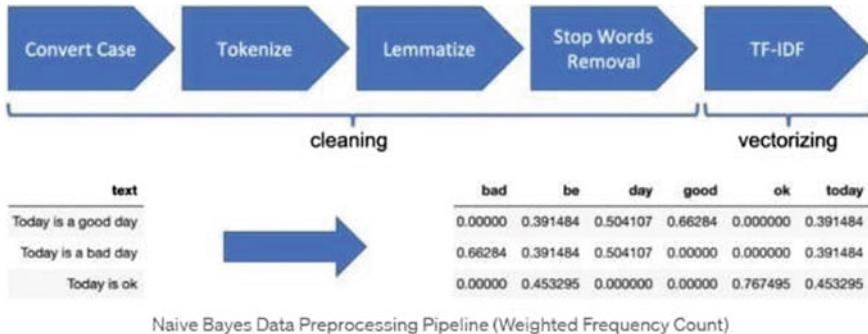
Next step is to clean the raw text corpus to mitigate any distortions in the text classifier filter. All the punctuations are removed for further process lemmatizer is used. For prediction a dictionary of words is created and various labels are assigned to map each word.

After this, the dataset is ready to be processed i.e., to perform different algorithms for sentiment analysis.

#### 4 Algorithm Analysis

**Naïve Bayes Algorithm:** The Naïve Bayes classifier makes assumptions according to previously learned potentially related conditions. In other words, the outcome is determined or concluded by the probabilities based on the condition of each lexical feature that is present in positive or negative form in our data being trained.

Machine Learning for NLP necessitates pre-processing in order to extract features from the original manuscript. Before the algorithms can process corpus data, it must be converted to numerical representations, a process known as vectorization [4]. We will create a simple DTM (document-term matrix) for Naive Bayes model consumption, but additional features such as text length, time/location of publication, named entities, and so on can be included.



**Fig. 6** Naïve Bayes data processing pipeline [1]

Because each distinctive word or phrase in the corpus's lexicon becomes a feature, DTM produces a significant featuring space. Some data cleaning steps are included to help with model (classifier) effectiveness and dimensionality reduction (Fig. 6).

### VADER (Valence Aware Dictionary and Sentiment Reasoner)

To capture context in the text, such as negations, N-grams (consecutive phrases) are employed. N-gram contexts are a blunt mechanism that might not or does not always correctly capture expressions. They might also have negative impact on the model by overburdening it with features.

The VADER lexicon is a dictionary that assigns a sentiment score to each characteristic, which might be a word, acronym, or emoticon, ranging from -4 (very negative) to 4 (extremely positive). VADER developed a language based on valence that can detect sentiment strength as well as polarity.

The input text's scores are computed using this, as well as powerful modifiers including negations, contractions, conjunctions, booster words, degree adverbs, capitalization, punctuation, and slangs.

On the test data set, we were able to acquire a 72% accuracy score using VADER. VADER considers emoticons and acronyms/initializations that are common in social media content, therefore it can do well in some domains but struggle in others. VADER generates accurate polarity scores, which we categorise as positive or negative for comparison purposes.

**Long-Short Term Memory Classifier:** Long-Short Term Memory (LSTM) is an abbreviation for Long-Short Term Memory. In terms of memory, LSTM is a type of recurrent neural network that outperforms traditional recurrent neural networks [8, 9]. LSTMs perform significantly better in the cases of memorising specific patterns. Recurrent neural networks called Long Short-Term Memory (LSTM) networks can learn order dependence in sequence prediction challenges. In complicated problem domains like machine translation, speech recognition, and others, this is a required behaviour. LSTMs are a difficult subset of deep learning.

When used with an LSTM layer, the Bidirectional wrapper propagates the input forward and backward through the LSTM layer and then concatenates the outputs.

Approach	Custom Machine Learning Models		Lexicon & Rule –Based Tool
	Naïve Bayes DTM	Word Embedding + LSTM Deep Learning	VADER
Training Effort	Model is fast to train	Resource intensive and takes time	Pre-trained models are easy and quick to implement
Interpretability	Outputs are reasonably understandable	Difficult to interpret model outputs	Results are easy to understand
Strength	Custom models capture context of the text are useful for domain-specific NLP tasks		<ul style="list-style-type: none"> <li>• Less resource and computationally intensive</li> <li>• Doesn't suffer severely from a speed-performance tradeoff</li> <li>• VADER contains linguistic rules that go beyond what is captured in a typical document-term-matrix model.</li> </ul>
	<ul style="list-style-type: none"> <li>• Computational is fast</li> <li>• Widely used for large-scale sentiment analysis</li> <li>• Retains context to a degree</li> </ul>	<ul style="list-style-type: none"> <li>• Has the potential to produce the most accurate results</li> <li>• Retains context of the text corpus</li> <li>• Less effort for feature extraction when neural networks can learn important features</li> </ul>	
Weakness	Supervised Machine Learning requires sufficient training data and feature extraction		<ul style="list-style-type: none"> <li>• Lexicon is difficult to create and validate</li> <li>• Susceptible to misspellings, nomenclatures, sarcasm, irony, jargons, and grammatical mistakes as they are not recognized by Lexicon</li> <li>• Ignores context of text</li> <li>• Suitability issues across domains</li> </ul>
	<ul style="list-style-type: none"> <li>• Naïve Bayes models rely on complete and representative data set</li> <li>• Strong independence assumption in the model</li> </ul>	<ul style="list-style-type: none"> <li>• Most expensive to train and operate</li> </ul>	

**Fig. 7** Comparison of sentiment analysis algorithms [8]

This assists LSTM in learning long-term dependencies. We then used it to train a dense neural network for classification (Fig 7).

## 5 Conclusion

Sentiment analysis is highly used in review systems by organization for their products and services. In the proposed system the social media website has a classifier filter which work on comment section which takes place when enable, it will filter out profane comments. For this filter we have analyzed various algorithms such as Naïve Bayes, LSTM and VADER. Implementation of profanity filter classifier is done after scrutinizing and analyzing the accuracy of above discussed sentimental analysis algorithms. For the social media website its UI is designed over figmas templates.

## References

1. D.M. Hussein, A survey on sentiment analysis challenges. J. King Saud Univ. Eng. Sci. (2018)
2. A. Sood, M. Hooda, S. Dhir, M. Bhatia, An initiative to identify depression using sentiment analysis: a machine learning approach. Ind. J. Sci. Technol. **11**(4) (2018)

3. A.H. Alamoodi, B.B. Zaidan, A.A. Zaidan, O.S. Albahri, K.I. Mohammed, R.Q. Malik, E.M. Almahdi, M.A. Chyad, Z. Tareq, A.S. Albahri, H. Hameed, Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: a systematic review. *Exp. Syst. Appl.* (2021)
4. L. Yang, Y. Li, J. Wang, R.S. Sherratt, Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* (2020)
5. L.I. Biloussova, L.E. Gryzun, N.V. Zhytienova, Fundamentals of UI/UX design as a component of the pre-service specialist's curriculum, in *SHS Web of Conferences* (2021)
6. D.M. Selfa, M. Carrillo, M.D. Boone, A database and web application based on MVC architecture, in *16th International Conference on Electronics, Communications and Computers, CONIELECOMP'06* (2006)
7. H.K. Ra, H.J. Yoon, A. Salekin, J.H. Lee, J.A. Stankovic, S.H. Son, Software architecture for efficiently designing cloud applications using node.js, in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion* (2016)
8. M. Sundermeyer, R. Schlüter, H. Ney, LSTM neural networks for language modeling, in *Thirteenth Annual Conference of the International Speech Communication Association* (2012)
9. D.P. Baruah, S. Dhir, M. Hooda, Impact of IOT in current era, in *International Conference on Machine Learning, Big Data. Cloud and Parallel Computing (COMITCon)* (2019)

# Artificial Intelligence Based Techniques for Underwater Image Restoration and Enhancement



Yogesh Kumar Gupta and Khushboo Saxena

**Abstract** Under-water picture enhancement has received much attention in under-water vision research. Under-water picture suffers from serious color distortion and low contrast problems because of complex light propagation in the ocean. However, raw under-water pictures easily suffer from underexposure, and fuzz caused by the under-water scene. Many of the research are going on to recover this under-water blurred image. Artificial intelligence based machine learning (ML) and deep learning (DL) techniques are optimizing better restoration rather than the conventional approaches. This paper presents an artificial intelligence based techniques for underwater image restoration and enhancement.

**Keywords** Underwater · Artificial · Intelligence · Image · Restoration

## 1 Introduction

Digital picture processing is developing day by day acquainting novel advances with give help to a few applications like mechanical exercises, under-water net arrangement, etc. Specifically, under-water picture preparing is considered as the pivotal undertaking in picture handling industry. While picture reconstruction innovation can enough consider eliminating this equivalent murkiness from source pictures, they need to acquire a few pictures from a specific spot that keep it from being utilized in an ongoing system, [1, 2]. Because of the dispersing and weakening of light into the water, the under-water picture generally shows up with shading twisting, obscured subtleties, and low differentiation. To resolve these issues, a clever two-stage under-water picture convolutional neural net (CNN) in view of design decay (UWCNN-SD) for under-water picture improvement is proposed by thinking about the qualities of

---

Y. K. Gupta · K. Saxena (✉)

Department of Computer Science, Banasthali Vidyapith, Banasthali, India

e-mail: [khushboosaxena.research20@gmail.com](mailto:khushboosaxena.research20@gmail.com)

Y. K. Gupta

e-mail: [gyogesh@banasthali.in](mailto:gyogesh@banasthali.in)

under-water imaging. In particular, the crude under-water picture is decayed into high-recurrence and low-recurrence dependent on hypothetical investigation of the under-water imaging [3].

The high-recurrence part is improved straight by a profound learning net. In any case, they can't precisely communicate the inherent attributes of the pictures, presenting frail transformation capacity for certifiable complex situations. All the more as of late, profound learning strategies for picture dehazing have developed and turn out to be more dependable, showing extraordinary execution. By and by, these techniques intensely rely upon preparing information, confining their application ranges. All the more critically, both conventional and profound learning approaches all disregard a typical issue, clamors/antiques consistently show up in the recuperation cycle [4].

Figure 1 presents the (a) under-water blur picture (the original captured picture) and (b) under-water picture (after processing of the methodology) of the sample image. The low quality of optical imaging brought about by the mind boggling and differing underwater climate is a huge test to under-water objective acknowledgment. Shading remedy for under-water pictures has gotten expanding revenue, because of its basic job in tasking with accessible mature vision calculations for under-water situations. Sound system cameras are broadly utilized for detecting and route of under-water mechanical frameworks.

They can give high goal shading perspectives on a scene; the compelled camera math empowers metrically precise profundity assessment; they are additionally generally financially savvy. Conventional sound system vision calculations depend on include identification and coordinating to empower triangulation of focuses for assessing divergence.

From the outset, utilize the initial CNN to change over the information under-water picture into the grayscale picture. Then, upgraded the grayscale under-water picture constantly CNN. And afterward, we play out the shading revision to the info underwater picture by the third CNN. Finally, can get the shading adjusted picture by coordinating the yields of three CNNs dependent on the tint safeguarding. In our system, that CNNs practice on each task can have the option to task on every engineering of CNNs probably and further develop the relapse quality to accomplish the low figuring cost and high effectiveness [5].

## 2 Literature Survey

Ramkumar et al. [1] presents strategy to conquer the under-water reconstruction issue, a profound report approach is created by giving magnificent results of profound learning approaches in a few other picture investigation concerns. A convolution neural net (CNN) model is prepared to de-cloudiness the singular pictures with picture reconstruction to perform further with picture improvement by including a standard picture input and here, the neural net is assessed by utilizing pictures and

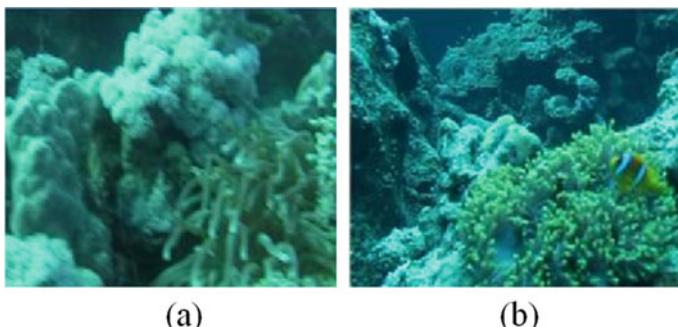
provisions, which are acquired from isolated regions to demonstrate its ability to sum up.

Wu et al. [3]. Presents, a two-stage under-water upgrade net including a primer improvement net and a refinement net is proposed. In the main stage, the primer improvement net, which contains the high-recurrence and the low-recurrence upgrade nets, is proposed. In the subsequent stage, the refinement net is intended to additionally streamline the shade of the under-water picture by thinking about intricacy of underwater imaging.

Liu et al. [4] gives, another Hadamard-Item (HP) model is proposed, which comprises of a progression of information driven priors. In light of this model, we infer a Learnable Hadamard-Item Spread (LHPP) by falling a progression of guideline propelled direction and recuperation modules. In which, the standard motivated direction identified with transmission is blessed the perfection property, the other recuperation module fulfills the dispersion of normal pictures. The Hadamard-item based proliferations are created in our created learnable structure for the assignment of picture dehazing. Along these lines, we can wipe out commotions/antiquities in the recuperation technique to acquire the best yields.

Pei et al. [6] shows the picture arrangement execution drops with every sort of debasement, regardless of whether this drop can be abstained from by including corrupted pictures into preparing, and whether existing PC vision calculations that endeavor to eliminate such corruptions can assist with further developing the picture characterization execution. In this article, we experimentally concentrate on those issues for nine sorts of corrupted pictures-murky pictures, movement obscured pictures, fish-eye pictures, under-water pictures, low goal pictures, salt-and-peppered pictures, pictures with white Gaussian commotion, Gaussian-obscured pictures, and out-of-center pictures.

Tune et al. [7] shows the strategy comprises of three significant stages. In the first place, the dataset with 84 pictures is increased (flip, adding commotion, and GAN (generative ill-disposed nets)) to 430 pictures, and all pictures are upgraded with MSRCR to task on their characteristics; Second, the model is pre-prepared on the COCO (Microsoft normal articles in setting) dataset to abbreviate the preparation



**Fig. 1** **a** Under-water blur picture, **b** under-water picture



**Fig. 2** Under-water pictures blur to restore

time and defeat overfitting; At long last, the pre-prepared model is moved to the under-water dataset, and the entire preparing measure is finished. We accomplish 97.46% accuracy and 94.52% review.

Chen et al. [8] presents a further developed picture super-goal recreation calculation dependent on profound convolutional neural net. The wavelet premise which can adequately recreate the waveform and qualities of under-water disturbance is chosen to supplant the neuron fitting capacity to task on the precision and productivity of the calculation. A further developed thick square design (IDB) is brought into the net which can viably tackle the slope vanishing issue of profound convolutional neural net and further develop the preparation speed simultaneously.

Liu et al. [9] shows, multiscale highlights are separated first, trailed by enlarging nearby components in each scale with worldwide elements. This plan was confirmed to task with more compelling and quicker net getting the hang of, bringing about better execution in both shading adjustment and detail protection. We directed broad investigations and contrasted the outcomes and best in class approaches quantitatively and subjectively, showing that our technique accomplishes critical enhancements.

Ueda et al. [10] proposes a technique to create blended under-water pictures from clean RGB-D pictures taken on the ground. It is valuable for preparing a profound neural net for under-water picture reconstruction (UWIR), and furthermore for estimating the exhibitions among UWIR strategies. The under-water pictures are incorporated on the demonstrating of a precise corruption measure with the thought of ingestion and dissipating just as ten water types.

Skinner et al. [11] shows, profound learning procedure applied to the issue of under-water picture reconstruction. Once more, it is troublesome or difficult to assemble genuine ground truth information for this issue. In this task, we present an unaided profound neural net (DNN) that takes input crude shading under-water sound system picture and yields thick profundity guides and shading remedied picture of under-water scenes. We influence a model of the course of under-water picture arrangement, picture preparing procedures, just as the mathematical requirements innate to the sound system vision issue to foster a particular net that beats existing strategies.

Figure 2 is showing the blur images and restores images of sample images. These pictures are taken to understand the blur and reconstructed images. The various blur image and processed images show in this figure.

Yang et al. [12] gives a locale identification net is proposed to gain proficiency with the connection between the cloudy picture and the medium transmission map in a patch wise way; the transmission map is then used to eliminate cloudiness by means of a climatic dispersing model and upgrade the detail of de-hazed pictures. The other is an original module with a fell cross channel pool, which wires staggered murkiness applicable elements and lifts the reflection capacity of the model on a nonlinear complex.

Barbosa et al. [13] presents a CNN-based methodology that doesn't need ground truth information since it utilizes a bunch of picture quality measurements to direct the reconstruction learning measure. The examinations showed that our strategy tasked on the visual nature of under-water pictures safeguarding their edges and furthermore performed well thinking about the UCIQE metric.

Protas et al. [14] presents, Convolutional Neural Nets (CNNs) have accomplished best in class execution in many picture reconstruction applications. The information on how these models task, in any case, is as yet restricted. In the task, present net reversal, another technique grew explicitly to help in the comprehension of picture reconstruction Convolutional Neural Nets. We apply our technique to under-water picture reconstruction and dehazing CNNs, showing how it can help in the arrangement and improvement of these models.

### 3 Challenges

After extensive research covering of previous years of research articles, it is strongly believe that under-water photography has confirmed its value in modern under-water photography, but comparing under-water photography with scientific and computerbased testing is a daunting task. Assistant requests. Other research posts are also identified as follows:

- At greater depths, under-water pictures are often less visible due to the narrowing of small streams.
- The presence of sea ice and small particles floating in the cooling of the camera and photographic aircraft leads to the distribution of gravity and rear, providing low light and blurry pictures.
- In addition, the light decreases with the lateral scattering distance, reducing the camera light and thus losing the natural colour.

Therefore, the main purpose of improving the under-water picture is to remove the black cover and colour compensation using the under-water picture. In proposed research, I will try to correct basic thinking, namely limited visibility, random lighting and diversity by overcoming the limitations and weaknesses of existing methods by building under-water picture enhancements and retrieval methods.

## 4 Proposed Strategy

The proposed methodology is based on the followings steps:

### (i) Collect data set

For execution the exploration errands, the submerged picture informational index will be taken structure the Kaggle ML correspondence or some other picture information supplier server.

### (ii) Preprocess of data

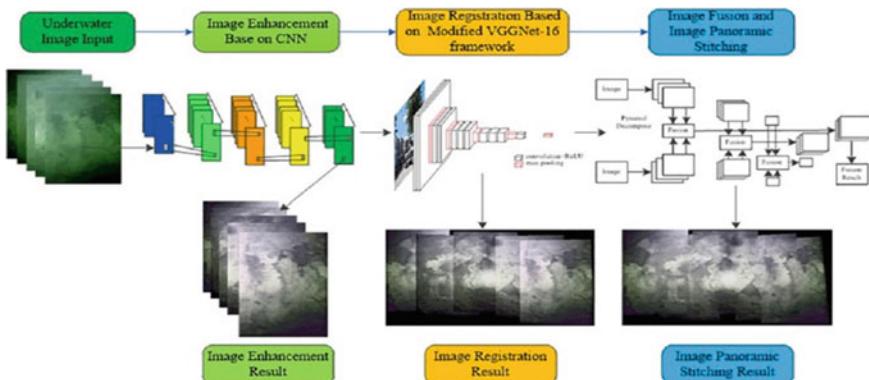
For each image, the accompanying preprocessing steps were applied:

1. Crop the piece of the image that contains just the submerged (which is the main piece of the image).
2. Resize the image to have a state of  $(240, 240, 3) = (\text{picture\_width}, \text{picture\_height}, \text{number of channels})$ : since pictures in the dataset come in various sizes. In this way, all photos ought to have a similar shape to take care of it as a contribution to the neural net.
3. Apply standardization: to scale pixel esteem to the reach 0–1.

### (iii) Model or system Model Architecture or configuration

A profound Convolution Neural Net (CNN) model with SVM is arranged that requires negligible pre-handling. The model can recognize and anticipate solid tissue, effectively upgrading cancer and non-propelling growth areas. The neighborhood invariant nature of CNNs takes into account deliberation of token highlights for arrangement without depending for enormous scope spatial data.

A symmetric convolution and deconvolution CNN-based structure for picture enhancement is introduced in Fig. 3. We constantly change the size of convolution portions and the quantity of element maps for streamlining the capacity of learning



**Fig. 3** CNN based model

degraded elements of the neural organization to acquire the nonlinear planning from submerged debased picture to improved picture. The haze debased submerged picture is taken as information and the size of the picture is limitless. The secret layer of introduced network is comprised of component maps, including three convolution subnets and deconvolution subnets with various convolution bits, and the two subnets are even. The convolution layer is performed to remove highlights, and the fluffy debased elements can be gained from the submerged corrupted pictures. Different element maps are created from convolution pieces. Because of the consistent convolution of CNN neglects to reestablish the subtleties of low quality submerged pictures, the even deconvolution layers are involved to refine the separated surface highlights, and it can reproduce the first picture utilizing the element.

#### (iv) **Mark removal**

A trait of these huge informational collections is an enormous number of factors that require a great deal of figuring assets to process. Mark removal is a course of dimensionality decrease by which an underlying arrangement of crude information is diminished to more sensible gatherings for handling.

#### (v) **Evaluation**

Three principle measurements used to assess an arrangement model are exactness, accuracy, and review. Precision is characterized as the level of right expectations for the test information. It tends to be determined effectively by partitioning the quantity of right expectations by the quantity of all out forecasts. DL is a subset of ML where fake neural nets, calculations roused by the human mind, gain from a lot of information. DL permits machines to take care of mind boggling issues in any event, when utilizing an informational index that is extremely different, unstructured and between associated.

#### (vi) **Simulation and Results**

Proposed technique will be carried out either MATLAB 9.4 form or Python Spyder IDE 3.7 variant. There will be upgrade in exactness and other huge boundaries utilizing proposed approach.

Presently apply the ML procedure to track down the presentation boundaries. The current assignment applied a few methods.

Presently the exhibition boundaries are determined as far as accuracy, review, f-1 measure, exactness and so forth by utilizing the accompanying recipes:

Correct Positive: anticipated valid and occasion are positive. Correct Negative: Predicted valid and occasion are negative.

Wrong Positive: anticipated wrong and occasion are positive. Wrong Negative: Predicted wrong and occasion are negative.

$$\text{Precision} = \frac{\text{CP}}{\text{CP} + \text{WP}}$$

$$\text{Recall} = \frac{\text{CP}}{\text{CP} + \text{WN}}$$

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Accuracy} = \frac{\text{CP} + \text{CN}}{\text{CP} + \text{CN} + \text{WP} + \text{WN}}$$

This formula will be used for calculation of the performance parameters by using proposed methodology implementation.

## 5 Conclusion

The picture handling local area has seen astounding advances in upgrading and reestablishing pictures. All things considered, reestablishing the visual nature of submerged pictures stays an extraordinary test. Start to finish structures may neglect to upgrade the visual nature of submerged pictures since in a few situations it isn't practical to give the ground reality of the scene brilliance. This paper study about the artificial intelligence based techniques for underwater image restoration and enhancement. The CNN based model will be used for future implementation of the proposed work.

## References

1. G. Ramkumar, A.G.S.K.M.M. Ayyadurai, C. Senthilkumar, An effectual underwater image enhancement using deep learning algorithm, in *5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1507–1511 (2021). <https://doi.org/10.1109/ICI-CCS51141.2021.9432116>
2. Y. Guo, H. Li, P. Zhuang, Underwater image enhancement using a multiscale dense generative adversarial net. In *IEEE J. Oceanic Eng.* **45**(3), 862–870 (2020). <https://doi.org/10.1109/OE.2019.2911447>
3. S. Wu et al., A two-stage underwater enhancement net based on structure decomposition and characteristics of underwater imaging. *IEEE J Oceanic Eng.* (2021). <https://doi.org/10.1109/JOE.2021.3064093>
4. Liu, S. Li, J. Liu, L. Ma, X. Fan, Z. Luo, Learning hadamard-product-propagation for image dehazing and beyond. *IEEE Trans Circuits Syst. Video Technol.* **31**(4), 1366–1379 (2021). <https://doi.org/10.1109/TCSVT.2020.3004854>
5. C. Yeh, C. Lin, C. Huang, Deep learning underwater image color correction enhancement based on hue preservation, in *IEEE Underwater Technology (UT)*, p. 16 (2019). <https://doi.org/10.1109/UT.2019.8734469>.
6. Y. Huang, Y. Pei, Q. Zou, X. Zhang, S. Wang, Effects of image degradation and degradation removal image classification. *IEEE Trans. Mach. Intell.* **43**(4), 1239–1253 (2021). <https://doi.org/10.1109/TPAMI.2019.2950923>

7. S. Song, J. Zhu, X. Li, Q. Huang, Integrate MSRCR and Mask R-CNN to recognize underwater creatures on small sample datasets. *IEEE Access* **8**, 172848172858 (2020). <https://doi.org/10.1109/ACCESS.2020.3025617>
8. Y. Chen, K. Niu, Z. Zeng, Y. Pan. A Wavelet based deep learning method for underwater image super resolution reconstruction. *IEEE Access* **8**, 117759–117769 (2020). <https://doi.org/10.1109/ACCESS.2020.3004141>
9. X. Liu, Z. Gao, B. M. Chen, MLFcGAN: multilevel feature fusion-based conditional GAN for underwater image color correction. *IEEE Geosci. Remote Sens. Lett.* **17**(9), 1488–1492 (2020). <https://doi.org/10.1109/LGRS.2019.2950056>
10. T. Ueda, K. Yamada, Y. Tanaka, Underwater image synthesis from RGB-D images and its application to deep underwater image reconstruction, in *IEEE International Conference on Image Processing (ICIP)*, pp. 2115–2119 (2019). <https://doi.org/10.1109/ICIP.2019.8803195>
11. K.A. Skinner, E.A. Olson, J. Zhang, M. Johnson-Roberson, UWStereoNet: unsupervised learning for depth estimation of underwater stereo imagery, in *International Conference on Robotics and Automation (ICRA)*, pp. 7947–7954 (2019). <https://doi.org/10.1109/ICRA.2019.8794272>
12. X. Yang, H. Li, Y.-L. Fan, R. Chen, Single image haze removal via region detection net. In *IEEE Trans. Multimedia* **21**(10), 2545–2560 (2019). <https://doi.org/10.1109/TMM.2019.2908375>
13. W.V. Barbosa, H.G.B. Amaral, T.L. Rocha, E.R. Nascimento, Visual-Quality driven learning for underwater vision enhancement, in *25th IEEE International Conference on Image Processing (ICIP)*, pp. 3933–3937 (2018). <https://doi.org/10.1109/ICIP.2018.8451356>
14. E. Protas, J. Bratti, J. Gaya, P. Drews, S. Botelho, understading image reconstruction convolutional neural nets with net inversion, in *16th IEEE International Conference on ML and Applications (ICMLA)*, pp. 215–220 (2017). <https://doi.org/10.1109/ICMLA.2017.0156>

# A Rapid Computing Technology on Profound Computing Era with Quantum Computing



Ramnarayan, Minakshi Memoria, Anuj Kumar, and Sunil Ghildiyal

**Abstract** The area of quantum computing bring the strong bridge of new computing era. For the automotive sector, quantum computing will have real-world ramifications, benefits, and concerns. It merely gives a cursory overview of current efforts to create quantum computers, a lively field that is still so new that even specialists cannot anticipate which approaches will be most effective. They define quantum computation topics precisely and underline important distinctions. There are many ways to explore the scientific laws in general and physics laws in particular. Knowledge, like physical laws of nature, can be conveyed in a diversification of ways. These ideas form the foundation for quantum algorithms, which use quantum computers to tackle complicated problems. This article explains some of the fundamental quantum physics concepts that underpin quantum computing. As they endeavor to outfit significantly a greater amount of the quantum world's latent capacity, there might be very nearly second quantum unrest. Many industries, including healthcare, energy, banking, security, and entertainment, could be affected by quantum computing and quantum communication. According to recent studies, the quantum sector will be worth billions of dollars by 2030.

**Keywords** Quantum computing · Quantum bits · ANN and quantum physics theory

## 1 Introduction

Quantum mechanics (QM) arose from a series of contentious mathematical explanations of experiments that classical mechanics could not explain. It began about the same time as Albert Einstein's theory of relativity, a separate mathematical revolution in physics that describes the motion of objects at high speeds. Unlike relativity, the beginnings of quantum mechanics cannot be traced back to a single scientist. Rather, between 1900 and 1930, a group of scientists worked together to provide

---

Ramnarayan (✉) · M. Memoria · A. Kumar · S. Ghildiyal  
UIT, Uttarakhand University, Dehradun, India  
e-mail: [ram000010@rediffmail.com](mailto:ram000010@rediffmail.com)

the groundwork for three innovative notions that eventually acquired recognition and experimental confirmation. This is wherein quantum PC frameworks get their aspect over traditional ones. In conditions wherein there is a tremendous scope of doable blends, quantum PC frameworks can review them at the same time. Models comprise of searching for the top components of an absolutely tremendous reach or the top notch heading among places. In any case, there can likewise be bunches of conditions wherein old style PC frameworks will regardless outflank quantum ones. So the PC frameworks of the predetermination can be a total of each those sorts [1].

## 2 Quantum Computing

Quantum processing is an excellent blend of quantum physical science, software engineering, and data hypothesis. The reason for this book is to make this intriguing examination region open to an expansive crowd. Specifically, we attempt to help peruse span the reasonable and notational hindrances that different quantum figuring from regular registering. One benefit of managing hypothesis, and not with the endeavors to construct quantum PCs, is that the measure of quantum material science and supporting math required is decreased. They can foster every one of the vital quantum mechanics inside the book; no past openness to quantum physical science is required. They give cautious and exact portrayals of crucial ideas, for example, quantum state spaces, quantum estimation, and snare—prior to covering the standard quantum calculations and other quantum data handling undertakings, for example, quantum key appropriation and quantum instant transportation [3]. Tackling AI issues utilizing the current situation with a quantum PC has its own restrictions. Quantum PCs are perplexing frameworks because of their requirement for ideal segregation from the climate. Absence of segregation results in the qubits of quantum PCs losing lucidness. This property of de-coherence of qubits is a difficult issue when planning a quantum calculation as a quantum circuit on a close term quantum PCs [1].

### 2.1 *Literature Survey*

Physicists and PC researchers such as Charles H. Bennet of the IBM Thomas J. Watson Research Center, Paul A. Benioff of the Argonne Public Laboratory in Illinois, David Deutch of the University of Oxford, and Richard P. Feynman of Caltech first investigated the possibility of a computational gadget based on quantum mechanics in the 1970s and mid 1980s. The idea came to mind as scholars were debating critical constraints of calculation. In 1982, Feynman was one of the few who attempted to seriously develop a new type of computer based on quantum physical science norms. He created a theoretical model to demonstrate how a quantum framework may be used to perform calculations and to highlight how such a machine could be used as a test system for real-world quantum physical science difficulties.

In other words, a physicist may use a quantum mechanical computer to do tests in quantum material science. Feynman went on to say that quantum computers can handle quantum mechanical challenges that are impossible to solve on a standard computer [2].

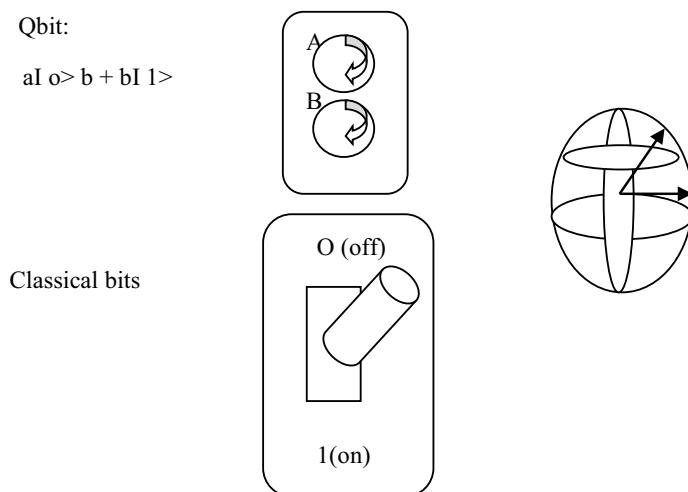
## 2.2 *Quantum Bits*

Customary PC utilizes two pieces 0 and 1 yet as they probably are aware quantum PC utilizes 0, 1 and mix of 0 and 1 all at once. Here 1 address the switch on of the semiconductor mean yes and the 0 address the off (no) condition. As per the Quantum physical science quantum mechanical simple piece is address as the quantum bit “Qbit”. The condition of the q-bits addressed numerically as (Fig. 1).

Execution of the Qbit is conceivable in various ways. A solitary number can be address in super conduction, zero energy or distinctive condition of processing levels [3].

### A. *Superposition*

The electron’s twist might be either in arrangement with the field, which is known as a twist up state, or inverse to the field, which is known as a twist down state. Changing the electron’s twist starting with one state then onto the next is accomplished by utilizing a beat of energy, for example, from a laser—suppose that we utilize 1 unit of laser energy. In any case, take under consideration the possibility that they just consume a large portion of a unit of laser energy and totally isolate the molecule from every outside influence. As indicated by quantum law, the molecule then, at



**Fig. 1** Qbit and classical bit representation

that point, enters a superposition of states, wherein maybe it was in the two states at the same time. Each qubit used could take a superposition of both 0 and 1. In this manner, the quantity of calculations that a quantum PC could attempt is  $2^n$ , where n is the quantity of qubits utilized. A quantum PC involved 500 qubits would possibly do  $2^{500}$  estimations in a solitary advance. This is a magnificent number— $2^{500}$  is vastly a greater number of iota than there are in the known universe (this is valid equal handling—old style PCs today, all things being equal called equal processors, still just genuinely do each thing in turn: there are only at least two of them doing it). Be that as it may, how might these particles communicate with one another? They would do as such through quantum trap [4].

### B. *Entanglement*

Particles that have cooperated eventually hold a sort of association and can be caught with one another two by two, in an interaction known as connection. Realizing the twist condition of one ensnared molecule—up or down—permits one to realize that the twist of its mate is the other way. Considerably more astounding is the information that, because of the peculiarity of superposition, the deliberate molecule has no single twist bearing prior to being estimated, yet is at the same time in both a twist up and turn down state. The twist condition of the molecule being estimated is chosen at the hour of estimation and conveyed to the related molecule, which all the while expects the contrary twist bearing to that of the deliberate molecule. This is a genuine oddity (Einstein dubbed it “scary activity a long distance off”), the instrument of which cannot be explained by any hypothesis at this time—it should essentially be considered as granted. Quantum ensnarement allows qubits separated by enormous distances to communicate with one another quickly (not restricted to the speed of light). Regardless of how incredible the distance between the corresponded particles, they will stay entrapped as long as they are separated [5].

## 3 Artificial Neural Network Approach to Quantum Computing

The goal of artificial neural organization in Quantum calculation is to limit the size of the PC components, which will be represented by the quantum laws. In 1950 the idea of artificial neural organization was presented. An artificial neural organization (ANN) over the ages, quantum processing has seen exceptional improvement which extraordinarily affects sped up registering. Identified with the counterfeit neural organization (ANN), an exceptional, advantageous, and pertinent idea has been suggested which is known as a quantum neural organization (QNN) [6]. QNN has been acknowledged joining the fundamentals of ANN with a quantum calculation standard which is better than the conventional ANN. Quantum PCs guarantee prominent benefits over old style PCs for a long time applications. A quantum computer employs some of quantum theory’ otherworldly miracles to achieve tremendous leaps ahead in processing capacity. Quantum machines confirmation to outperform even the most

ready to do today—and tomorrow—supercomputers. Man-made intelligence (ML), especially applied to significant neural associations through the backpropagation estimation, has permitted a wide extent of inventive applications loosening up from the social to the sensible.

### 3.1 *Intention*

To fabricate a totally quantum broad neural organization skilful in all-inclusive quantum calculation we have thought that it is essential to change the current recommendations. A quantum perceptron to be a typical unitary leader following up on the comparing information and yield qubits has been characterized by us, whose boundaries consolidate the loads and predispositions of past proposition impressively. Besides, we present a training calculation for this quantum neural organization that is powerful such that it just depends on the expansiveness of the singular layers and not on the profundity of the organization. We notice that the proposed network has some amazing properties, as the ability to sum up from tiny informational indexes and a striking defencelessness to uproarious preparing information [7].

## 4 Quantum Physics

Quantum physical science underlies how particles work, thus why science and science function as they do. You, I and the gatepost—at some level at any rate, all of us are getting into the quantum rhythm. Assuming one need to clarify how electrons travel through a central processor, how photons of light get gone to electrical flow in a sun powered charger or enhance themselves in a laser, or even exactly how the sun continues to consume, you will need to utilize quantum physical science. The trouble—and, for physicists, the fun—begins here. In any case, there is no single quantum hypothesis [8]. There is quantum mechanics, the essential numerical system that supports everything, which was first evolved during the 1920s by Niels Bohr, Werner Heisenberg, Erwin Schrödinger and others. It portrays straightforward things, for example, how the position or force of a solitary molecule or gathering of not many particles changes over the long haul. Be that as it may, to see how things work in reality, quantum mechanics should be joined with different components of material science—mainly, Albert Einstein's unique hypothesis of relativity, which clarifies what happens when things move exceptionally quick—to make what are known as quantum field speculations.

Three distinct quantum field hypotheses manage three of the four major powers by which matter communicates: electromagnetism, which clarifies how molecules hold together; the solid atomic power, which clarifies the solidness of the core at the core of the particle; and the frail atomic power, which clarifies why a few iota go through radioactive rot.

## 4.1 Working and Performance

In quantum processing, activities rather utilize the quantum condition of an item to deliver what is known as a qubit. These states are the vague properties of an item before they have been distinguished, like the twist of an electron or the polarization of a photon. Rather than having an unmistakable position, unmeasured quantum states happen in a blended 'superposition', much the same as a coin turning through the air before it lands in your grasp. These superposition can be entrapped with those of different articles, which mean their ultimate results will be numerically related regardless of whether we know yet what they are [8].

The mind-boggling science behind these disrupted situations of ensnared 'spinning coins' can be linked to unique computations that can quickly solve problems that would take an old-style PC a long time to figure out... assuming they could ever figure them out at all. Solving complex numerical problems, establishing difficult-to-crack security codes, and anticipating diverse molecular cooperation's in material responses will all be benefit from such computations.

## 5 Conclusion and Future Aspects

Beginning around the year 2020, quantum figuring has been a focus for the government. "We are approaching a time when traditional figuring power development, using the traditional method for silicon and semiconductor power, is attracting to a close proximity. We won't see the next wave of processing power development come from a combination of programming, new designs, and a general framework and worldview makeover" Minister of State (MoS) for Electronics and IT Rajeev Chandrashekhar stated [9]. "That is where Quantum Computing enters the picture, and it is clear that it will be at the forefront of future requests for registering power. QSim is a stepping stone for Indian researchers to get us on that route" Added he. The National Mission on Quantum Technologies and Applications, announced by Finance Minister Nirmala Sitharaman, would cost 8000 crore over five years. The Ministry of Electronics and Information Technology (MeitY) had previously announced collaboration with Amazon Web Services (AWS) to establish a Quantum Computing Applications Lab in the country, with the goal of giving designers, academics, and others access to quantum registering development conditions.

## References

1. S. Arunachalam, J. Briet, C. Palazuelos, Quantum query algorithms are completely bounded forms. *SIAM J. Comput.*, **48**(3), 903–925 (2019). Earlier version in ITCS'18. [arXiv:1711.07285](https://arxiv.org/abs/1711.07285)
2. P.A. Benioff, Quantum mechanical HAMILTONIAN models of turing machines. *J. Stat Phys.* **29**(3), 515–546 (1982)

3. J.-F. Biasse, F. Song, Efficient quantum algorithms for computing class groups and Solving the principal ideal problem in arbitrary degree number fields, in *Proceedings of 27th ACM-SIAM SODA*, pp. 893–902 (2016)
4. E. Bernstein, U. Vazirani, Quantum complexity theory. SIAM J. Comput. **26**(5), 1411–1473 (1997). Earlier version in STOC’93
5. A. Bookatz, QMA-complete problems. Quantum Inf. Comput. **14**(5–6), 361–383 (2014). [arXiv:1212.6312](#)
6. M. Boyer, G. Brassard, P. Hoyer, A. Tapp, Tight bounds on quantum searching. Fortschritte der Physik, **46**(4–5), 493–505 (1998). Earlier version in Physcomp’96. quantph/9605034
7. G. Brassard, P. Hoyer, M. Mosca, A. Tapp, Quantum amplitude amplification and estimation, in *Quantum Computation and Quantum Information: A Millennium Volume volume 305 of AMS Contemporary Mathematics Series*, pp. 53–74 (2002). quant-ph/0005055
8. Brassard, G., P. Hoyer, A. Tapp, Quantum algorithm for the collision problem. ACM SIGACT News Crypt Column **28**, 14–19 (1997). quant-ph/9705002
9. A. Broadbent, C. Schaffner, Quantum cryptography beyond quantum key distribution. Des. Codes Crypt. **78**(1), 351–382 (2016). [arXiv:1510.06120](#)
10. S. Aaronson, A. Ambainis, Quantum search of spatial regions. Theory Comput. **1**(1), 47–79 (2005). Earlier version in FOCS’03. quant-ph/0303041
11. R. Beals, H. Buhrman, R. Cleve, M. Mosca, R. de Wolf, Quantum lower bounds by polynomials. J. ACM **48**(4), 778–797 (2001). Earlier version in FOCS’98. quantph/9802049

# Covid-19: Comparison of Time Series Forecasting Models and Hybrid ARIMA-ANN



N. Hema Priya, S. M. Adithya Harish, N. Ravi Subramanian, and B. Surendiran

**Abstract** The COVID-19 outbreak has taken many countless lives in the society and presents an unprecedented public health threat, as well as a threat to food systems and workplace safety. As a result, various machine learning models that can forecast the outbreak internationally have been developed. In order to help in the restriction of the transmission and growing number of covid cases in India, this study employs a range of methodologies to forecast the total number of current cases in India over the following 15 days. To forecast the future, the ARIMA Model, Facebook Prophet and Holt's Winter Model are utilised. Before being preprocessed, data is gathered in real time from a variety of sources. After that, the data set is split into two sections: training and testing. Finally, the model's accuracy is trained and evaluated. The approaches' forecasting effectiveness is influenced by strategies for deconstructing the original data and merging linear and nonlinear models during the hybridization process, according to the results. By applying correct techniques, a hybrid approach can be an useful approach for enhancing forecasting accuracy achieved by traditional hybrid approaches as well as any of the other constituent methods employed alone. This paper provides a hybrid methodology that incorporates both ARIMA and ANN models to take advantage of the unique characteristics of ARIMA and ANN models in linear and nonlinear modelling. The Hybrid model showed better accuracy and a root mean square error of 21,267 which is the lowest among other models compared. The article emphasizes on the ability to assist governments in acting and making sound decisions, as well as planning for the future, in order to reduce public concern and prepare people's thinking for the next phases of the pandemic.

**Keywords** ARIMA · Hybrid ARIMA-ANN · Holt's winter model · Time series forecasting

---

N. Hema Priya · S. M. Adithya Harish (✉) · N. Ravi Subramanian  
Department of Information Technology, PSG College of Technology, Coimbatore, India  
e-mail: [adithyaharish10@gmail.com](mailto:adithyaharish10@gmail.com)

B. Surendiran  
Department of Computer Science and Engineering, National Institute of Technology, Puducherry, India

## 1 Introduction

Covid-19, the ongoing pandemic, has wreaked havoc on the global economy, causing governments to take drastic steps to stem its spread. As a result, many countries have introduced self-isolation and social structure isolating steps to prevent additional spread, straightening the pandemic curve, which could be crucial in ensuring that health facilities are offered to those who need them most, whether for COVID-19 or other serious conditions. In terms of infrastructure, finance, business, manufacturing, and a variety of other resources, our society is making enormous efforts to resist the spread of this terrible disease [1]. To be able to reliably predict when an outbreak will peak will help governments better their policies and plan for preventive measures such as general public wellness texts, enhancing citizen awareness, and increasing the capacity associated with wellness systems. As a result, multiple machine learning models that could predict the global pandemic have already been produced. This paper provides a quick examination and analysis of the key models that are used to anticipate COVID19. Time series forecasting is when you make scientific predictions based on previously time stamped data. It entails creating models based on historical data and using them to draw assumptions and steer essential choices in the future. A lot of effort has gone into building and enhancing time series forecasting models in recent years.

In this paper, we present a hybrid technique for simulating the linear and nonlinear correlation structures of a time series that incorporates ARIMA and ANN. We answer the question of how well a hybrid model performs in comparison to predicting models. The proposed hybrid method is based on the fact that real-world time series frequently contain both linear and nonlinear characteristics [2].

## 2 Related Works

This section describes the scientific work that went into this study. This contains both (1) previous studies that used machine learning and statistical methodologies to anticipate or foresee an event, and (2) more current studies that used similar approaches but solely predicted COVID-19 pandemic-concerned statistics like daily cases and mortalities. To solve the issue of time series forecasting using a Long Short-Term Memory (LSTM) neural network, Helli et al. [3] developed a Holt–Winters Additive Model (HWAAS). The LSTM neural network is a special sort of RNN can learn long-term relationships. The LSTM Network, which forecasted the total number of COVID-19 instances using an exponential linear unit, fared better. The paper “COVID-19: a comparison of time series techniques to forecast percentage of active cases per population,” by Vasilis Papastefanopoulos et al., looks at an automatic forecasting procedure that acts as a solution. It used time as a regressor to fit a variety of linear and nonlinear time functions as components. The method used a time-series model that can be broken down into three parts: trend, seasonality, and

vacations. The Prophet employed the Fourier series to compensate for seasonality. As a response, a more adaptable periodic impacts model was developed. In order to make up for previous and foreseeable vacations, this model required a stated list of the same [4, 5].

Dhamodharavadhani and Rathipriya [6] used a Gaussian Process to assign a probability to the infinite functions that could match the data. This expresses the model's uncertainty, and it's a good indicator of how much we can trust the point forecast. The Gaussian model has a lot of predictive potential, so it's utilised to anticipate the peak number of infected persons and when it will happen. With a root mean square error very low, the model fits the actual data well (RMSE). Though it's simple to control and employ, Gaussian processes have no unique characteristics that aren't replicated by other approaches, and the only drawback is computation time. Hybrid models were also employed by other researchers to increase forecast accuracy. In this work, linear and nonlinear models were merged. Combining linear and nonlinear models overcomes the constraints of employing just one type of strategy and produces more accurate results [7–9]. Some researchers have sought to alter data by combining pre- and post-processing approaches, as well as hybrid methods that combine linear and nonlinear models [10–12]. As a result, the hybrid models' forecasting abilities combined with data preparation and data cleaning operations can increase performance.

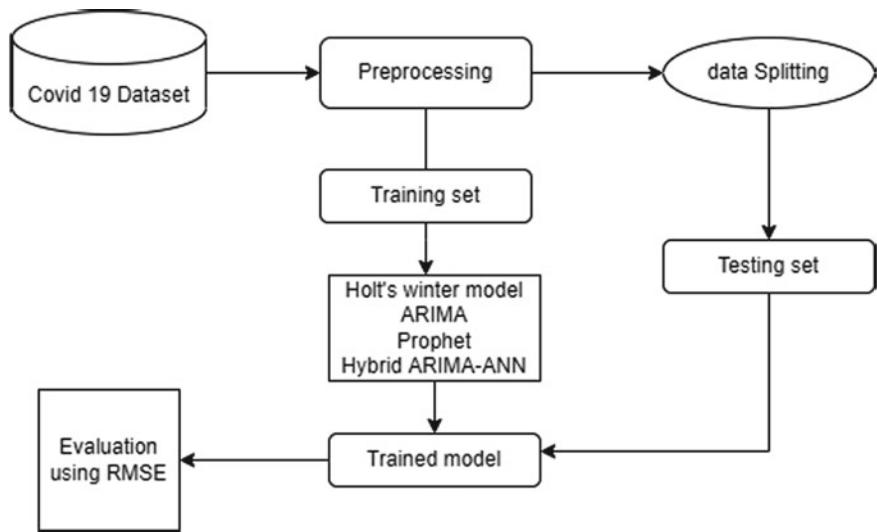
Earlier research in the related works section attempted to create a model that could accurately predict or foretell. In their methodology, this research attempted to anticipate using linear or nonlinear approaches. According to the findings of the literature review, hybrid models that incorporate both the ARIMA and the ANN can give better outcomes. As a result, a hybrid analytical framework combining ARIMA and ANN would be presented in the following section.

## 2.1 Dataset

In any prediction model, dataset is the key factor deciding efficiency as the attributes are what being trained. The data set used in this research is gathered from Kaggle [13] which has information on daily level information on the number of affected cases, deaths and recovery from 2019 novel coronavirus across the globe. The dataset consists of 306 k rows with 8 columns. We take the parameters Observation Date, Country/Region, Confirmed cases, Death cases into account for our evaluation.

## 3 Proposed Methodology

See Fig. 1.



**Fig. 1** Workflow of the model

### 3.1 Time Series Forecasting

Time series forecasting is a method for predicting events over a period of time. It forecasts future occurrences by examining historical patterns and assuming that future trends will follow in the footsteps of past patterns. A historical sequence is used to begin time series forecasting. Researchers look at historical data for periodic decomposition characteristics such as trends, seasonal changes, cyclic trends, and uniformity. ARIMA [14], the Holt–Winters additive model (HWAAS), and Facebook’s Prophet [15] are three forecasting systems that are depicted and briefly explained in this section.

### 3.2 Prophet Model

Prophet is a time series forecasting technology developed by Facebook to handle commercial time series concerns. A decomposable time-series model [15, 16] has three basic components: trend, seasonality, and holidays. Prophet forecasts trends using a saturating growth model and a piece-wise linear model. A model comparable to population increase models in ecological systems [17], in which nonlinear growth hits a critical threshold at carrying capacity, is employed to estimate growth. A piece-wise version of steady growth gives a useful and frequently valuable strategy for anticipating concerns when this saturation barrier is never achieved. Prophet [17] employs Fourier series to produce a scalable system of periodic effects, but

to correct for holidays, it requires an existing database of previous and prospective holiday happenings. Because vacation impacts are assumed to be self-contained, they are simple to incorporate into the model [15]. The mathematical representation is as follows:

$$y(t) = tr(t) + se(t) + hd(t) + r \quad (1)$$

- tr(t) trend function for modelling non-periodic changes in the statistics.
- se(t) periodic changes (yearly, weekly and daily).
- hd(t) function for impacts of holidays which can be provided by user.
- r model's noise/distraction.

### **3.3 ARIMA Model**

The Auto-Regressive Integrated Moving Average (ARIMA) is a type of time-series model that is well-known and commonly used. ARIMA models assume a linear relation between time-series values and use these linear relations in observations to uncover local structure while removing high-frequency distraction from the data [18]. This type of technique has two significant advantages. To begin with, it gives a high level of understandability because the link between the independent and dependent variables is well-understood and hence easily addressed based on the model's premises. ARIMA models may also govern systems with changing trends over time by upgrading the model to forecast the system's future outcome based on recent occurrences [19]. ARIMA models, but at the other end, fail to manage nonlinear patterns or relationships, hence they are not necessarily realistic in replicating complicated real-world issues and dynamics [20].

The value of  $Y_t$  in the AR model is determined by its own lagged value. The value of  $Y_t$  in the MA model is determined by lagged forecast errors. As a result, ARIMA's general equation is

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + r \quad (2)$$

where,  $Y_t$  is the series' lag1,  $\beta$  is the model's estimate of the lag1 coefficient and  $\alpha$  is the model's estimate of the intercept term and  $r$  is errors.

### **3.4 Holt's Winter Model**

The technique of repeatedly revising a forecast after keeping in mind more latest observations is known as exponential smoothing. In effect, this is accomplished by reducing the weights of older observations, which reduces their value for forecasting rapidly. In other words, more recent findings count more than older ones

for projecting a new value. For producing predictions, this method consists of three essential components. Because of the pattern and seasonality, it has a mean amount. The hold winter's approach is also called as triple exponential smoothing since the three components are three different types of exponential smoothing. Simple exponential smoothing is used for predicting when there are no trends or seasonality in the data collection. Holt's smoothing technique, also known as linear exponential smoothing, is a famous smoothing approach for forecasting data with a trend. As a result, the Holt winter's method includes the average as well as trend and seasonality when projecting time series.

$$\text{Forecast equation} = yt + h|t = \ell + hbt \quad (3)$$

$$\text{Level equation} = \ell = \alpha yt + (1 - \alpha)(\ell t - 1 + bt - 1) \quad (4)$$

$$\text{Trend equation} = b = \beta_*(\ell t - \ell t - 1) + (1 - \beta_*)bt - 1 \quad (5)$$

where,

$\ell$  is an estimation of the series level at time  $t$ ,  $b$  is a trend's estimation of the series at time  $t$ ,  $\alpha$  is the coefficient of smoothing.

### 3.5 Hybrid ARIMA-ANN

Machine learning methods have been proven to be effective for tackling time series problems in several studies, yet no single forecasting model seems to be applicable in all scenarios. Though the machine learning model outperforms traditional methods, solving all-time series problems with a single machine learning model would be challenging and impossible. To improve prediction accuracy, combining linear and nonlinear patterns is recommended, and Zhang's hybrid technique was used [14].

$$Y = L + N \quad (6)$$

where  $L$  denotes the linear component and  $N$  denotes the non-linear component. The nonlinear relationship will be obtained from the linear model's residuals. The linear component is initially fitted with ARIMA, and the associated forecast  $L_1$  at time  $t$  is obtained. Let  $e_{time}$  be the linear model's residuals at time  $t$ , then

$$e_{time} = Y_1 - L_1 \quad (7)$$

where  $L_1$  is the forecast value of time  $t$ . The ANN model for residuals with  $n$  input nodes will be

$$e_{time} = f(e_{time-1}, e_{time-2}, \dots, e_{time-n}) + r_t; \quad (8)$$

The neural network model determines  $f$ , a nonlinear function, and  $r_t$  is the random distraction. If  $N_1$  is this ANN's forecast, then the prediction from the Hybrid  $Y_1$  model can be composed as

$$Y_1 = L_1 + N_1 \quad (9)$$

Finally, the hybrid system technique that is advised is divided into two components. To begin, the linear element of the problem is examined using an ARIMA model. The ARIMA model's returns are then simulated using a neural network model. The residuals of the ARIMA model will contain information on the nonlinearity of the data because the ARIMA model cannot account for it. The ARIMA model's error terms can be forecasted using the neural network's outputs. To determine a wide range of patterns, the hybrid model incorporates the features and strengths of both the ARIMA and ANN models. ARIMA models may not be adequate for approximating challenging nonlinear situations. Using ANNs to represent linear challenges, on the other hand, has shown mixed results. It may be helpful to model linear and nonlinear trends separately using distinct models and then integrate the estimates to improve modelling and forecasting effectiveness.

## 4 Results

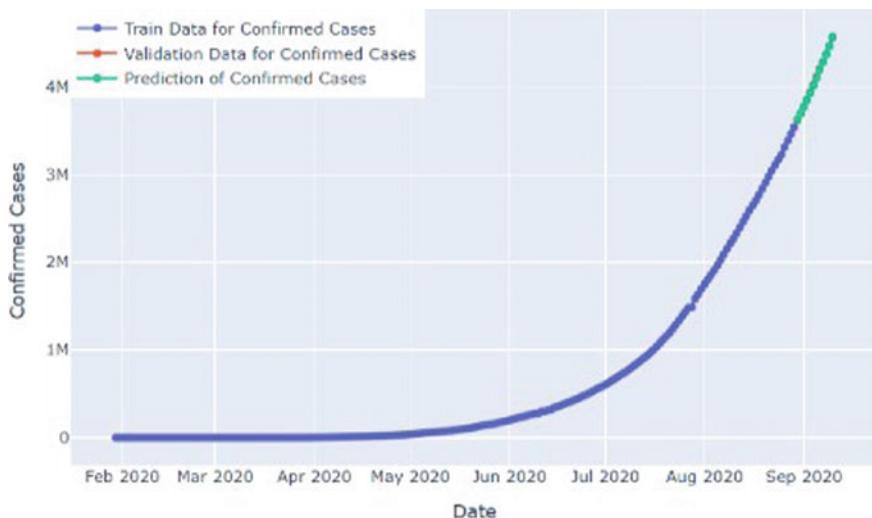
### 4.1 Evaluation Metric

The performance of learning models in terms of RMS error is examined in this study (RMSE). The RMSE (root of the error terms) is a measure of how far the extrapolation curve deviates from the actual data points [18, 21]. It's a metric that measures how the residuals spread out around the best-fit line. Because the units are identical to the output units, it will be simple to read. A smaller RMSE value improves model performance

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \quad (10)$$

### 4.2 Visualization of Models

After calculating the root mean square error, we plot the prediction graphically for better visualization. The closer trained and predicted values, better is the accuracy (Figs. 2 and 3).



**Fig. 2** Holt's winter model

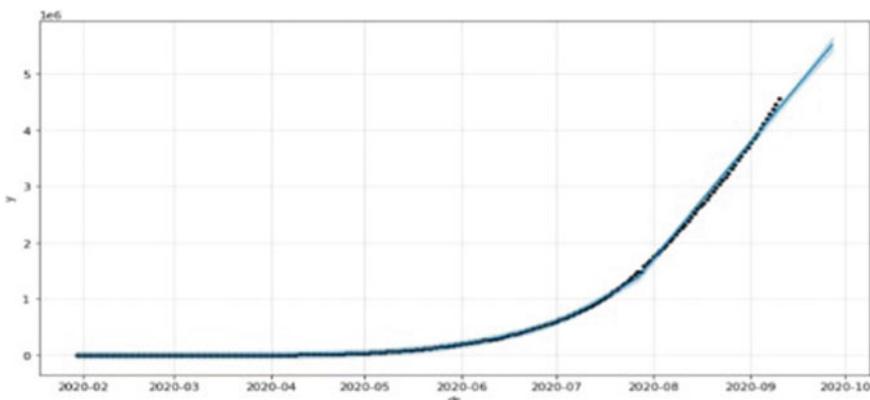
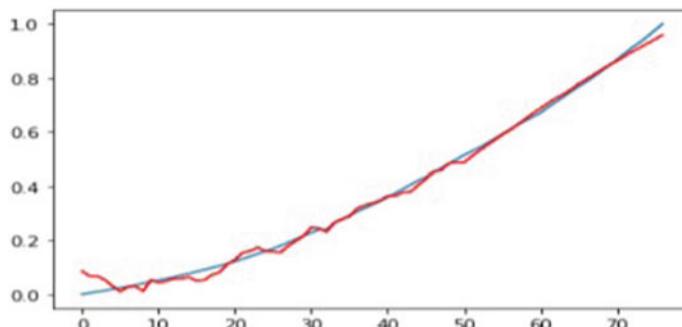


**Fig. 3** ARIMA model

The holt and ARIMA models are both quite efficient, as shown in the Table 1. We also propose a novel hybrid technique to boost accuracy. We suggest utilising a combining technique to forecast time series. The ARIMA and ANN models are used to integrate a variety of interactions in time series data. In linear and nonlinear modelling, the hybrid model combines the benefits of ARIMA and ANN. For complicated situations with both linear and nonlinear correlation patterns, the combination technique can be a valuable way to increase predicting performance (Figs. 4 and 5).

**Table 1** Comparison of forecasting models

Model name	RMSE	Mean absolute error
Facebook's prophet model	31,954.6	2601.15
ARIMA model	27,223.7	185.59
Holt's winter model	22,385.752	137.62
Hybrid ARIMA-ANN	21,067.269	135.72

**Fig. 4** Facebook prophet model**Fig. 5** Hybrid ARIMA-ANN model

## 5 Conclusion

In recent decades, time series analysis and forecasting have been a prominent focus of research. Despite the fact that ARIMA and ANNs are capable of dealing with a wide range of difficulties, neither is the ultimate chosen model for all forecasting situations. For nonlinearly generated time series, ARIMA is better, and for linearly generated time series, ANN is better. Because real-world time series frequently contain

both linear and nonlinear correlation trends, defining a series' exact nature is nearly difficult. We present a time series forecasting technique based on a combination of techniques in this research. To capture a range of interactions in time series data, the ARIMA model and the ANN model are merged. In linear and nonlinear modelling, the hybrid model combines the strengths of ARIMA and ANN. The combination technique can be a useful way to boost prediction power for complex tasks with both linear and nonlinear correlation trends. The hybrid model outperforms each component model when employed separately, with an RMSE of roughly 20 k, the lowest of any model, according to the experimental results from data sets.

## References

1. B. McCloskey et al., Mass gathering events and reducing further global spread of COVID-19: A political and public health dilemma. *The Lancet* **395** (10230), 10961099 (2020)
2. G.P. Zhang, Times series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175 (2003)
3. S.S. Helli et al., Short-term forecasting COVID-19 cases in turkey using long short-term memory network, in *2020 Medical Technologies Congress (TIPTEKNO)* (IEEE, 2020)
4. V. Papastefanopoulos, P. Linardatos, S. Kotsiantis, COVID-19: a comparison of time series methods to forecast percentage of active cases per population. *Appl. Sci.* **10**(11), 3880 (2020)
5. S. Singh, et al., Time series analysis of COVID-19 data to study the effect of lockdown and unlock in India. *J. Inst. Eng. (India): Series B* (2021)
6. S. Dhamodharavardhani, R. Rathipriya, J.M. Chatterjee, COVID-19 mortality rate prediction for India using statistical neural network models. *Front. Public Health* **8** (2020)
7. F. Ardakani, M. Ardehali, Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. *Energy* **65**, 452–461 (2014)
8. U. Yolcu, E. Egrioglu, C.H. Aladag, A new linear & nonlinear artificial neural network model for time series forecasting. *Decis. Support Syst.* **54**(3), 1340–1347 (2013)
9. A.S. Ahmad et al., A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renew. Sustain. Energy Rev.* **33** (2014), 102–109
10. C. Yuan, S. Liu, Z. Fang, Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. *Energy* **100**, 384–390 (2016)
11. T. Xiong, Y. Bao, Z. Hu, Does restraining end effect matter in EMD based modeling framework for time series prediction? Some experimental evidences. *Neurocomputing* **123**, 174–184 (2014)
12. Lin at al, Empirical mode decomposition-based least squares support vector regression for foreign exchange rate forecasting. *Econ. Model.* **29**(6), 2583–2590 (2012)
13. <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>
14. G.E.P. Box et al., *Time Series Analysis: Forecasting and Control* (Wiley, 2015)
15. S.J. Taylor et al., Forecasting at scale. *Am. Stat.* **72** (1), 37–45 (2018)
16. A.C. Harvey, S. Peters, Estimation procedures for structural time series models. *J. Forecast.* **9**(2), 89–108 (1990)
17. A.C. Harvey, Estimation and testing of stochastic variance models. No. 268. Suntory and Toyota International Centres for Economics and Related Disciplines LSE (1993)
18. C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **30**(1), 79–82 (2005)
19. M.J. Kane et al., Comparison of ARIMA and random forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinform.* **15** (1), 1–9 (2014)

20. H. Wold, A study in the analysis of stationary time series. Diss. Almqvist & Wiksell (1938)
21. W. Wang, Y. Lu., Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model, in *IOP Conference Series: Materials Science and Engineering*, vol. 324, no. 1 (IOP Publishing, 2018)

# YOLOV5 and Morphological Hat Transform Based Traffic Sign Recognition



Ponduri Vasanthi and Laavanya Mohan

**Abstract** This article narrates the traffic sign recognition by using morphological hat transform and YOLOV5. Recognition of traffic signs while driving on the road is necessary to estimate traffic situations and to avoid the accident. Due to bad weather conditions and night time, it is difficult to detect traffic signs in low contrast images. To tackle this problem, in this article, morphological hat transform used to enhance low contrast images and YOLOV5 used for detection. For this traffic sign detection, YOLOV5 network is trained with 4 classes of traffic sign dataset, which contains totally 740 images, in these 592 images used for training and 148 images used as validation images. This YOLOV5 network detected traffic signs with precision 77.9, recall 93.0 and obtained 0.78 mAP.

**Keywords** Traffic sign recognition · Object detection · Deep learning · YOLOV5 · Bounding box

## 1 Introduction

Now a days, traffic sign detection and estimating of the traffic situation are necessary, while driving the vehicle. If driver miss out the road sign detection, then it prone to dangerous situations. Sometimes, driving is difficult for driver when roads are unfamiliar. At that situations traffic sign recognition plays an enormous role for getting road safety and for ease of driving. In this article, traffic sign recognition is achieved by using YOLOV5 algorithm.

Object detection is a part of computer vision. Object detection is a task, it can detect multiple objects and identifies the location of the objects which are present in an image. Convolutional neural networks [1] plays amendment role in object detection. object detection is achieved by using two approaches. One is two-step approach and second one is one-step approach. In two-step approach, selected features are extracted from the image and convolution is applied for detection task. R-CNN [2] family

---

P. Vasanthi (✉) · L. Mohan

Vignan's Foundation for Science, Technology and Research, Guntur, Andhra Pradesh, India  
e-mail: [vasanthi457@gmail.com](mailto:vasanthi457@gmail.com)

coming under two-step approach. In one-step approach, convolution directly applied on the entire input image. SSD and YOLO family come under the one step-approach [3]. YOLO versions gives excellent results in object detection. YOLOV2 gives better accuracy and speed compare with the YOLOV1. YOLOV3 [4, 5] gives higher accuracy compare with YOLOV2 but speed is less than the YOLOV2. YOLOV4 has high mAP compare with the YOLOV3 [6]. But YOLOV4 train the network more slowly. YOLOV5 overcome all these existed models' drawbacks and gives high accuracy [7].

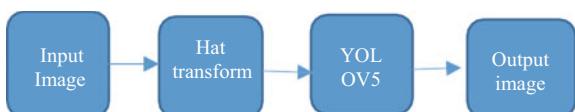
To perceive traffic signs some scholars introduced some approaches such as Sun et al. [8] planned a model to perceive and categorize the circular signs. Image pre-processing is used to highlight the essential information then areas are located by using Hough transform. Hough transform detected areas fed to CNN, it categorizes the traffic signs. Djebbara Yasmina et al. introduced modified LeNet-5, improved CNN and data augmentation [9]. In this, modified LeNet-5 network is used to carry out deep features from the input image. Convolutional layers output is connected to the multilayer perceptron (MLP) in improved CNN and finally data augmentation techniques are used to increase the accuracy. Alexander Shustanova and Pavel Yakimova planned GHT (Generalised Hough transform) algorithm with modifications [10]. In this, the generalised Hough transform (GHT) algorithm has been changed for focusing the road sign then classification done by using CNN. Marco Magdy William planned Faster RCNN and SSD with mobile Net V1, ResNet, Inception V2 and tiny YOLO v2 [11]. In which, various detection models are compared. Faster R-CNN and SSD with mobile Net V1, Inception V2 and tiny YOLO v2 are used to detect the traffic signs while the presence of various weather situations such as illumination and visibility challenges. Jyoti parsola et al. proposed an automated system, it is used for extraction of road signs and guessing of traffic volume [12].

## 2 Proposed Method

Apart from detection, the traffic sign recognition has suffered with many problems due to weather conditions, such as unable to yield abundant information, edges description is not clear, low contrast and significant features of road sign images are not permissible into the detection result. This article is going to introduce an enhanced technique by using morphological hat transformation.

Block diagram of introduced algorithm has been shown in Fig. 1. First source images size are  $415 \times 415$  images. Enhancement of source images is done by using morphological hat transformation. Later YOLOV5 algorithm is used to detect traffic

**Fig. 1** Block diagram of proposed method



signs. After YOLOV5 detection, output images consist of traffic signs along with labels and confidence scores.

The top-hat transform is a digital image processing technique that extracts minor features and details from input images. The difference between the input image and its opening by some structural element is described by the white top-hat transform, and the difference between the closure and the input picture is described by the black tophat transform. Feature extraction, background equalization, picture enhancement, and other image processing activities are all done via top-hat transformations [13]. The names are Open Top-Hat or White Top-Hat and Close Top-Hat or Black Top-Hat transformations. Many studies use the term Top-Hat to refer to both types of hat transformations [14]. The mathematical equations for top-hat and bottom-hat can be written as,

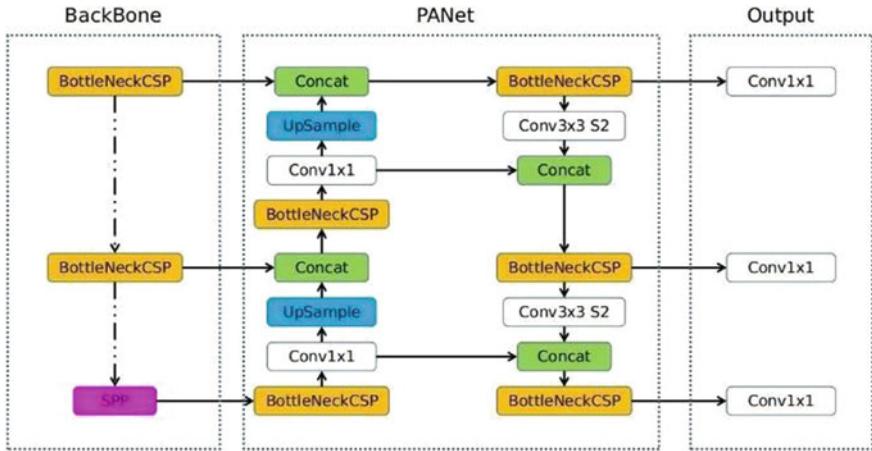
$$\text{Top-hat} = A - (A \circ B) \quad (1)$$

$$\text{Bottom hat} = (A * B) - A \quad (2)$$

YOLOV5 architecture consists of three stages i.e., first stage is model backbone, second stage is Model neck and third stage is Model head. The feature is extracted from the input image using the backbone stage. In YOLOV5, a backbone network called a cross stage partial network is utilized to extract informative characteristics from the input image. Feature pyramids are generated by model necks, these are helpful for identify the similar objects with different scales and sizes. Available feature pyramid models are FPN, BIFPN and PANET etc. In YOLOV5, PANET used to obtain feature pyramids. The function of model head is the detection. Model head generates anchor boxes on objects and generates output class probabilities, scores along with labels. In the hidden layers of the network, YOLOV5 employs a leaky RELU activation function, whereas the final detection layer employs a sigmoid activation function. The SGD optimization approach is used by YOLOV5 (Fig. 2).

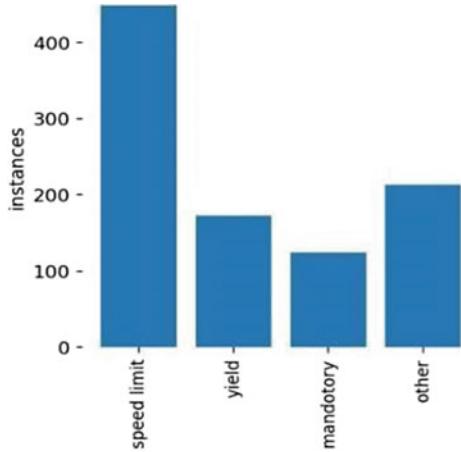
### 3 Dataset

The traffic sign recognition dataset taken from the Kaggle website [15]. This dataset consists of 740 images with 4 classes in which training images are 592 and validation images are 148. Figure 3 shows the instance of classes in which speed limit sign images are more, mandatory sign images are very low compare with the yield and other.



**Fig. 2** YOLOv5 architecture

**Fig. 3** Instances of classes



## 4 Results

There are some traffic sign images, which are shown in Fig. 4, in column 1, that are hard or impossible to analyse and extract information from input images, because of their low contrast. As a result, we should seek for innovative ways to improve the quality of such input images. The suggested solution, is mathematical morphological hat transform, it improves picture contrast and heals scratched areas. In fact, white features are extracted using a Top-Hat transformation, enhanced images are shown in Fig. 4, in column 2.

YOLOv5 network is trained with image size of  $415 \times 415$  and batch size of 16 for 30 epochs. After training the following are the training results are precision, recall,



**Fig. 4** Row 1: input images, row 2: enhanced images

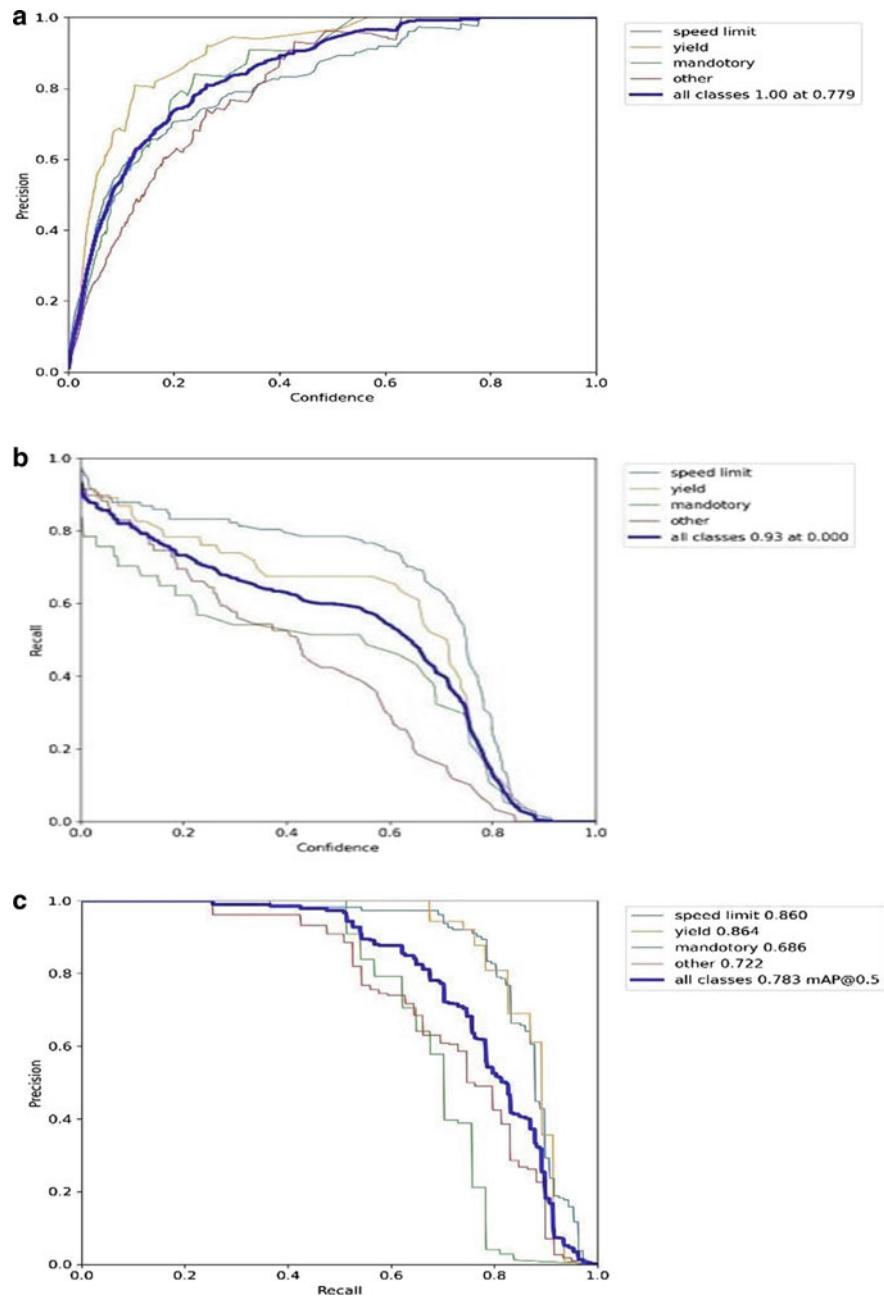
precision and recall, F1 curve, train loss, validation loss and mAP metrics are shown in Fig. 5a–d (Table 1).

All four classes i.e. speed limit, yield, mandatory and other gives precision = 1 at confidence level = 0.779 which is shown in Fig. 5a. All four classes i.e. speed limit, yield, mandatory and other gives recall = 0.93 at confidence level = 0.00 which is shown in Fig. 5b. All four classes i.e. speed limit, yield, mandatory and other gives F1 score = 0.93 at confidence level = 0.262 which is shown in Fig. 5d.

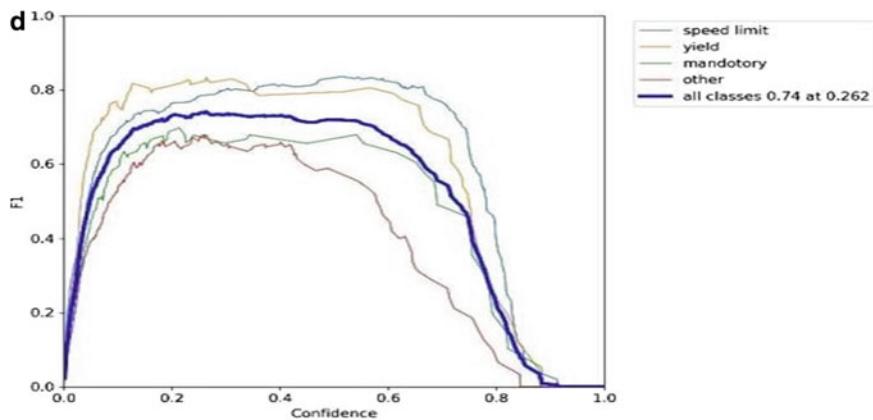
For 30 epochs, Fig. 6 shows train/box\_loss, train/cls\_loss, train/obj\_loss, val/box\_loss, val/cls\_loss, val/obj\_loss, metrics/precision, metrics/recall, metrics/mAP 0.5, and metrics/mAP 0.5–0.95. Increasing the number of epochs reduces train/box\_loss, train/cls\_loss, train/obj\_loss, val/box\_loss, val/cls\_loss, and val/obj\_loss. The number of epochs increases the metrics/precision, metrics/recall, metrics/mAP 0.5, and metrics/mAP 0.5–0.95 (Table 2).

Predicted class labels and their confidence scores are shown in Fig. 7. Totally sixteen images are shown in Fig. 7. In this, four classes of images detected with labels and detection probabilities. After applying test image to the model, YOLOV5 detected speed limit traffic sign with 0.57 score and other sign with 0.31 score are shown in Fig. 8 as well as in Table 3.

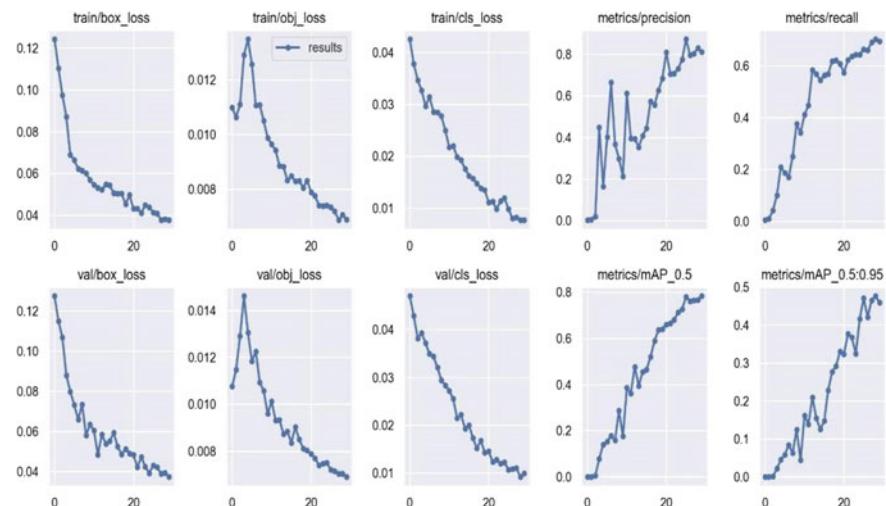
From Table 4, our suggested model is better to all existing deep learning approaches in boosting the contrast and quality of the input images to increase the accuracy of the detection process because it goes through an exfoliation process. Unlike other deep learning systems, this method's image resolution is extremely high, and image details are better represented. Finally, traffic indicators are quickly noticed.



**Fig. 5** **a** Precision curve. **b** Recall curve. **c** Precision and recall curve. **d** F1 curve

**Fig. 5** (continued)**Table 1** P-R curve value for different classes

S. No.	Class	P-R curve value
1	Speed limit	0.860
2	Yield	0.864
3	Mandatory	0.686
4	Other	0.722
5	All classes	0.783 at mAP@0.5

**Fig. 6** Train loss, validation loss and mAP metrics

**Table 2** Train and validation losses, precision, recall and mAP for 30 epochs

Epoch	train/box_loss	train/obj_loss	train/cls_loss	Metrics/precision	Metrics/recall	Metrics/mAP_0.5	Metrics/mAP_0.5:0.95	val/box_loss	val/obj_loss	val/cls_loss
30	0.037	0.006	0.007	0.779	0.93	0.78	0.45	0.037	0.006	0.009



**Fig. 7** Class predicted image



**Fig. 8** Output test image

**Table 3** Detected objects with labels and detection probability

S. No.	Class	Detection probability
1	Speed limit	0.57
4	Other	0.31

**Table 4** Comparision of state of art models with hat transform based YOLOV5

S. No.	Model	Precision (%)	Recall (%)	mAP_0.5 (%)
1	Faster RCNN	74.12	87.95	85.84
2	R-FCN	76.69	89.16	87.19
3	SSD	70.41	76.05	75.11
4	RetinaNet	69.83	75.71	75.02
5	YOLOV3	70.16	77.4	76.92
6	YOLOV4	69.71	78.5	78.05
7	YOLOV5	71.92	80.31	80.05
8	TS-YOLO	74.53	84.01	83.73
9	Hat transform based YOLOV5	77.9	93.0	78.3

## 5 Conclusion

In this article, to improve the quality and contrast of traffic sign photographs, morphological Top-Hat transforms are applied. Later YOLOV5 algorithm applied on traffic sign data, for detection of traffic signs. This hat transform based YOLOV5 model detects objects very accurately with precision of 77.9, recall of 93.0 and mAP of 78.3. Hence, traffic signs are detected very precisely by using YOLOV5 with confidence scores compared with the other deep learning models. Because of its advantages, YOLOV5 can be implemented in different fields to tackle some real-life problems like road security, monitoring traffic lanes and avoid the road accident.

## References

1. L. Mohan, V. Veeramani, Image denoising with a convolution neural network using Gaussian filtered residuals. *IEIE Trans. Smart Process. Comput.* **10**(2), 96–100 (2021)
2. V. Vijayaraghavan, M. Laavanya, Vehicle classification and detection using deep learning. *Int. J. Eng. Adv. Technol.* **9**, 24–28 (2019)
3. Z.Q. Zhao, P. Zheng, S.T. Xu, X. Wu, Object detection with deep learning: a review. *IEEE Trans. Neural Netw. Learn. Syst.* **30**(11), 3212–3232 (2019)
4. W. Li, K. Liu, L. Zhang, F. Cheng, Object detection based on an adaptive attention mechanism. *Sci. Rep.* **10**(1), 1–13 (2020)
5. V. Ponduri, L. Mohan, Reliable multiple object detection on noisy images by using YOLOV3. *Int. J. Commun. Comput. Technol.*, 6–9 (2021). <https://doi.org/10.31838/ijcts/09.01.02>
6. X. Zhang, W. Wang, Y. Zhao, H. Xie, An improved YOLOV3 model based on skipping connections and spatial pyramid pooling. *Syst. Sci. Control Eng.*, 1–8 (2020)
7. Y. Yin, H. Li, W. Fu, Faster-YOLO: an accurate and faster object detection method. *Digit. Sig. Process.* **102**, 102756 (2020)
8. Y. Sun, P. Ge, D. Liu, Traffic sign detection and recognition based on convolutional neural network, in *Chinese Automation Congress (CAC)* (IEEE, 2019), pp. 2851–2854
9. D. Yasmina, R. Karima, A. Ouahiba, Traffic signs recognition with deep learning, in *2018 International Conference on Applied Smart Systems (ICASS)* (IEEE, 2018), pp. 1–5

10. A. Shustanov, P. Yakimov, CNN design for real-time traffic sign recognition. Proc. Eng. **201**, 718–725 (2017)
11. P.S. Zaki, M.M. William, B.K. Soliman, K.G. Alexsan, K. Khalil, M. El-Moursy, Traffic signs detection an recognition system using deep learning. arXiv preprint [arXiv:2003.03256](https://arxiv.org/abs/2003.03256) (2020)
12. J. Parsola, D. Gangodkar, A. Mittal, Automated system for road extraction and traffic volume estimation for traffic jam detection. Int. J. Comput. Vis. Robot. **11**(2), 127–150 (2021)
13. H. Hass nanopour, N. Samadiani, S.M. Salehi, Using nsforms to enhance the contrast of medical images. Egypt. J. Radiol. Nucl. Med. **46**(2), 481489 (2015)
14. S.K. Panguluri, L. Mohan, An enhanced image fusion framework using morphological operations based unsharp masking, in *2021 International Conference on Computer Communication and Informatics (ICCCI)* (IEEE, 2021), pp. 1–6
15. Traffic-signs-dataset. <https://www.kaggle.com/valentynsichkar/traffic-signs-dataset-in-yolo-format>

# Optimizer Based Performance Study of Block Wise Fine-Tuned (BWFT) AlexNet Architecture



Simran Kaur Hora and Maya Ingle

**Abstract** Convolutional Neural Networks (CNN) is frequently utilized in computer vision, image classification, and other applications. Due to shape similarities, classification and recognition of handwritten Devanagari characters has been a challenging task. In our study, transfer learning is proposed to fine-tune AlexNet using a handwritten Devanagari character dataset to evaluate the performance of network block wise using various optimizers. A BWFT approach is used for classifying around 39,600 images from Devanagari dataset and implemented in Matlab. The study has been presented using three learning rates and optimizers as Adam, Sgdm, Rmsprop along with the performance of block-wise training of CNN architecture AlexNet as it is the focus of this study. AlexNet is fine-tuned using Devanagari datasets after being pre-trained on the ImageNet dataset. Finally, we conclude that when AlexNet network is fine-tuned with a Learning Rate (LR) of  $10^{-4}$ , it performs best for recognition purpose on Adam and Rmsprop optimizers and hence is highly advisable in developing AI applications.

**Keywords** Convolutional neural network · Image classification · Transfer learning · Fine-tuning

## 1 Introduction

Automatic classification and recognition of Devanagari script has been of wide interest in recent research, owing to its importance in a number of applications. Some examples of applications include digital signature verification, automatic bank check processing, automatic handwritten text detection in classroom teaching etc. [1–3]. In Devanagari handwritten character recognition, feature extraction approaches such as Discrete Cosine Transformation (DCT), Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT) etc. and classifiers such as neural networks, KNN (k-Nearest Neighbors), Naïve Bayes (NB), SVM (Support Vector Machine) [4,

---

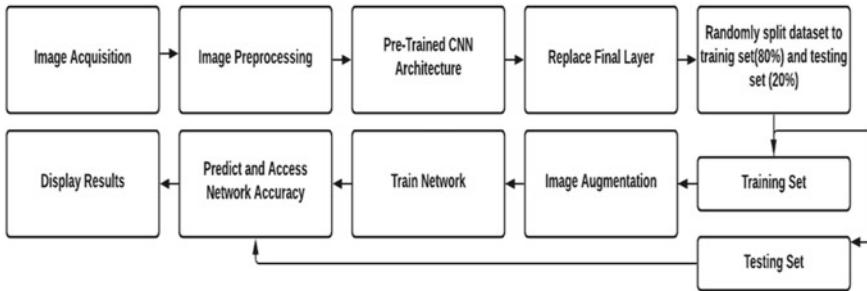
S. K. Hora (✉) · M. Ingle  
SCS&IT, DAVV, Indore, India  
e-mail: [simranhora701@gmail.com](mailto:simranhora701@gmail.com)

[5] are commonly employed. An accuracy of 90.70% is achieved using DCT feature extractor with SVM classifier [4], whereas utilizing Gabor feature extractor with SVM classifier attains an accuracy of 91.39% [5]. Attempts are made for classification and recognition of Devanagari handwritten numerals and characters. Convolutional Neural Network (CNN) has gained popularity tremendously in domains such as image classification, pattern identification etc. as these are capable of extracting feature and classify images in one step. Recently, it has been observed that there are numerous powerful pretrained architectures of CNN such as VGGNet [6], ResNet [7], AlexNet [8], and are used widely in AI application domains. To train the CNN from scratch and to overcome overfitting issues, it usually requires a significant quantity of data. Transfer Learning (TL) may be utilized in two ways; by involving either Feature Extraction techniques or Fine-Tuning. In feature extraction, network designed for one task is used for the next task whereas in fine-tuning, a pre-trained network is utilized as a starting point for learning new tasks [9]. Transfer Learning involving fine-tuning AlexNet achieved an accuracy ranging between 93.8 and 95.7% using parameters as initial LR:  $3e^{-4}$ , mini batch size:128 for 100 epochs [10]. Fine-tuning AlexNet with parameters as LR:  $1e^{-4}$  and min batch: 128 gained an accuracy of 99.39% [11]. The importance of CNN in research community has grappled due to AI applications in practice. Using CNN for Devanagari character recognition a maximum recognition accuracy of 93.73% is attained using 75% training data and 25% test data at 30 epochs [12]. However, there is a scope of studying the approach that may competitively be proven to be not only an alternative but also efficient for Devanagari characters classification.

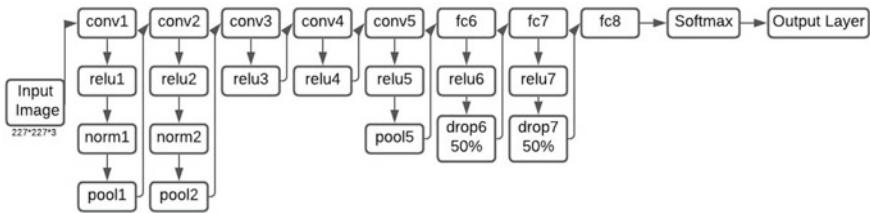
In this paper, a study of BWFT AlexNet using various optimizers based on performance parameters as learning rates and accuracy for the Devanagari dataset is proposed and elaborated. Section 2 discusses the approach used for the study of BWFT in detail. The experimental execution using the approach based BWFT and the respective results are emphasized in Sect. 3. At the end, we conclude in Sect. 4.

## 2 Approach Used for the Study

Our Block Wise Fine Tuned (BWFT) AlexNet approach consists of various phases namely; Image acquisition, Image Pre-processing, Pre-Trained CNN Architecture, Replacement of Final Layer, Randomly Splitting of dataset, Image Augmentation, Train Network, Predict Network Accuracy and Display Results as shown in Fig. 1. This approach is firstly used for classification of Devanagari characters in Devanagari dataset. In Image acquisition, the input images (in.png form at) are acquired from a publically accessible dataset available at UCI machine learning repository. Image Preprocessing is responsible to scale the images so as to meet the requirement of input size as per the architecture of AlexNet. Pre-trained CNN Architecture focuses on fine-tuning AlexNet using transfer learning. AlexNet architecture is comprised of five convolutional layers, maxpooling layers, three fully connected layers and a liner layer with softmax activation at the output as depicted in Fig. 2. A regularization



**Fig. 1** BWFT AlexNet approach for image classification



**Fig. 2** Pre-trained CNN architecture

procedure known as dropout is utilized for reducing overfitting in fully connected layers. The ReLU activation function is implemented on each of the first seven layers. Furthermore, the input size of images must be adhere the size as  $227 \times 227$  pixels as per the input pixel size criterion of AlexNet [13]. The image characteristics extracted by convolutional layers are employed by the final classification layer to classify the input image. To retrain a pre-trained network to classify Devanagari images, replace the last fully-connected layer to a new fully connected layer with the same number of outputs equal as the number of classes in Devanagari dataset. Dataset Splitting is concerned with partitioning of dataset randomly into two divisions for training and validation. Image Augmentation inhibits the network from overfitting and remembering the specific information included in the training images. Image augmentation is a method of creating artificial images by altering the underlying images using transformations such as scaling, rotations, and other adjustments [14]. In Train Network, the network is trained using options such as mini batch size, learning rate, epoch size, optimizers. At the stage of Predict and Access Network Accuracy, validation images are categorized using BWFT. It helps to compute the classification accuracy in order to predict and test the accuracy of the fine-tuned network. Display Result deals with the display of images with their predicted labels and probabilities on the screen.

### 3 Experimental Execution and Results

To study the performance of BWFT AlexNet approach on different optimizers, a specific setup has been used and is described below.

#### 3.1 *Experimental Setup and Execution*

For this study, Matlab on Windows 10 as the software platform is used. The datasets namely; ImageNet is employed to train CNN and Devanagari Handwritten Character Dataset to fine-tune the AlexNet. ImageNet offers millions of photos that have been manually labeled for our study [15] whereas Devanagari Handwritten Character Dataset is an image database consisting of 36 different annotated characters [16]. A total of 39,600 images of characters have been extracted manually for further experimentation.

Our BWFT AlexNet approach is employed and executed on a mini-batch size of 64. For maintaining the consistency, epoch size is fixed as 10. The experiment has been performed using three LR viz.  $10^{-4}$ ,  $10^{-5}$ , and  $10^{-6}$ . The optimizers used here for performance study are namely; Adaptive Moment Estimation (Adam), Stochastic Gradient Descent with Momentum (Sgdm), Root Mean Square Propagation Algorithm (RMSProp). Sgdm uses a single learning rate for all the parameters. Rmsprop seek to improve network training by using learning rates that differ by parameter and can automatically adapt to the loss function being optimized. Adam uses a parameter update that is similar to RMSProp, but with an added momentum term [17]. The input images Devanagari dataset are scaled and randomly partitioned into 80 and 20% for training and validation respectively. The library imageDataAugmenter of Matlab is preferred for all the image augmentation operations in training the data. The backward fine tuning of AlexNet is achieved using deep learning toolbox.

#### 3.2 *Experimental Results*

Our BWFT AlexNet approach is executed for Devanagari image dataset for fixed epoch size, using different LR and optimizers as mentioned earlier. It has been observed that validation accuracy is attained as 91.53% by fine-tuning blocks from 5th to 2nd on Sgdm optimizer whereas accuracy of 94.15% is achieved by fine-tuning blocks from 5th to 2nd on Rmsprop optimizer. On the other hand, it has been noticed that the performance of BWFT attains the validation accuracy as 94.19% by finetuning blocks from 5th to 1st on Adam optimizer as shown in Table 1. Using the function subplot of Matlab, predicted labels are displayed along with the probabilities of validated images by BWFT AlexNet. Some random validated images on Adam and corresponding probabilities have been shown in Fig. 3.

**Table 1** Performance of fine-tuned AlexNet on different LRs and optimizers

Block No.	Accuracy (in %)									
	Sgdm			Rmsprop			Adam			
	LR			LR			LR			
	$10^{-4}$	$10^{-5}$	$10^{-6}$	$10^{-4}$	$10^{-5}$	$10^{-6}$	$10^{-4}$	$10^{-5}$	$10^{-6}$	
5th	86.63	79.58	58.35	91.89	89.70	79.68	89.82	89.14	80.86	
4th	88.62	81.99	57.49	91.62	91.30	80.42	92.31	90.33	80.92	
3rd	90.77	82.59	56.93	92.46	92.99	82.56	91.26	91.62	82.40	
2nd	91.53	83.65	57.97	94.15	92.12	83.98	91.70	91.99	85	
1st	91.36	84.52	57.18	91.54	92.07	84.73	94.19	92.75	84.61	
All	91.49	83.96	59.12	93.78	92.93	85.09	92.69	92.53	85.93	

**Fig. 3** Predicted labels and probabilities by BWFT

## 4 Conclusion

The performance of our BWFT AlexNet based on Devanagari dataset using various optimizers is the focus of this study. Using three optimizers *Sgdm*, *Adam* and *Rmsprop* the comparative performance of BWFT AlexNet using Devanagari datasets is presented. It is to conclude that when BWFT AlexNet is fine-tuned with a Learning Rate of  $10^{-4}$ , it attains the highest accuracy as 94.19% for recognition of Devanagari characters on Adam optimizer. Thus, the impact of our results may be useful for AI application developers.

## References

1. T.A. Pham, H.H. Le, N.T. Do, Offline handwritten signature verification using local and global features. *Ann. Math. Artif. Intell.* **75**, 231–247 (2014)
2. R. Palacios, A. Gupta, A system for processing handwritten bank checks automatically. *Image Vis. Comput.* **26**(10), 1297–1313 (2008)
3. P. Banerjee et al., Automatic detection of handwritten texts from video frames of lectures, in *14th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (2014), pp. 627–632
4. S.R. Narang et al., Devanagari ancient character recognition using HOG and DCT features, in *Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (2018), pp. 215–220
5. S.R. Narang et al. On the recognition of Devanagari ancient handwritten characters using SIFT and gabor features. *Soft Comput.* **24**, 17279–17289 (2020)
6. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition (2015)
7. K. He, X. Zhang et al., Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 770–778
8. A. Krizhevsky et al., Imagenet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2012)
9. A. Brodzicki, M. Piekarski et al., Transfer learning methods as a new approach in computer vision tasks with small datasets. *Found. Comput. Decis. Sci.* **45**, 179–193 (2020)
10. A. Titoriya, S. Sachdeva, Breast cancer histopathology image classification using AlexNet, in *2019 4th International Conference on Information Systems and Computer Networks (ISCON)* (2019), pp. 708–712
11. M.E. Morocho-Cayamcela et al., *Fine-Tuning a Pre-trained Convolutional Neural Network Model to Translate American Sign Language in Real-Time* (2019), pp.100–104
12. S.R. Narang, M. Kumar et al., DeepNetDevanagari: a deep learning model for Devanagari ancient character recognition. *Multimedia Tools Appl.* **80**, 20671–20686 (2021)
13. K. Zhang, Q. Wu et al., Can deep learning identify Tomato leaf disease? *Adv. Multimedia* (2018)
14. I. Kandel, M. Castelli, How deeply to fine-tune a convolutional neural network: a case study using a histopathology. *Appl. Sci.* (2020)
15. J. Deng, W. Dong et al., ImageNet: a large-scale hierarchical image database, in *IEEE Conference on Computer Vision and Pattern Recognition* (2009), pp. 248–255
16. S. Acharya et al., Deep learning based large scale handwritten Devanagari character recognition, in *9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, 1–6 (2015)
17. S.R. Dubey et al., diffGrad: an optimization method for convolutional neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **31**, 4500–4511 (2020)

# Comparative Study on Image Feature Extraction Techniques



**Yogita Bhajgotar and Vaibhav Gandhi**

**Abstract** In image processing, algorithms are used to identify various desired parts or features (Shape, Edge) of the image. Using image processing we can detect shape of the object and extract texture of the objects. There are many methods for performing feature extraction in image processing depends on a domain. This paper discusses the comparison of the different image feature extraction techniques with parameters like image translation, scaling, rotation invariant, noise etc. The motivation is to obtain the best technique to be used in further feature extraction from images. The extracted features are chosen carefully; the feature set should extract the relevant information from the input data to perform the desired task. Feature extraction is about to obtain the relevant information from the original data. It will be useful in image classification and recognition, object detection. Extracting features like edges and corner from an image is useful for many purposes.

**Keywords** Image recognition · Image extraction · Image processing · Feature extraction

## 1 Introduction

Here we compare some common methods for feature extraction in image processing. Histogram of an image can also be used as a picture element, but this is a very naive technique. Wave-based techniques are also widely used. It is pixel processing algorithms that are used to locate points on an image that can be recorded with similar points on other images. Color Histograms represent the distribution of colors in an image. Here introduced two widely used shape features: Fourier descriptors and Zernike moments. Extracting features such as edges and curves from an image

---

Y. Bhajgotar (✉)  
Computer Engineering, GTU, Ahmedabad, Gujarat, India  
e-mail: [bhajgotaryogita@gmail.com](mailto:bhajgotaryogita@gmail.com)

V. Gandhi  
Department of MCA, B H Gardi College of Engineering and Technology, Rajkot, Gujarat, India

is useful for many purposes. There are many different edge detection methods. The most powerful edge detection method provided by edges is the canny method. Along with Canny and Prewitt, Sobel is one of the most popular edge detection algorithms used in technology today.

This paper discusses the details of the different image feature extraction techniques in Sects. 2, and 3 contains, comparison on various image feature extraction techniques: Color, edge, and shape respectively and comparison on the accuracy, rotation invariant, spatial information, noise, the computation cost of these methods is given. Finally, the conclusion is presented in last section.

## 2 Image Feature Extraction

Image feature provides rich information of image content such as edge, region, and interested point of the image. Features extraction is important technique in image processing used for object recognition and image classification [1].

### 2.1 Shape Feature

Shape features like triangular objects, circular object, and other shapes. There are two methods for feature extraction region-based and edge-based feature extraction methods [1–3].

#### Zernike Moments

Zernike Moment is widely used for shape feature. It is used to extract features from grayscale image. It is geometric based that is 2 dimensional function of orthogonal. It contains set of infinite complex polynomials of unit circle. Zernike Moments are invariant to rotation, scaling, translation, and insensitive to noise. It is suitable for pattern recognition and image reconstruction [4].

#### Geometric Moments

Image moments are used to describe objects after they have been segmented. It is used in texture classification, edge detection and pattern recognition. It generates scaling, translation and rotation invariant [5, 6].

#### Fourier Descriptor

Fourier Descriptors (FD) uses Fourier transformed boundary. It is robust features for objects representation and boundaries. 2 D boundary represented by 1 D transform and a point is  $(X, Y)$  is reduced by complex number  $X + jY$ . It is used to recognize 2 dimensional shapes. It is invariant to changes in translation, scaling, and rotation [6].

## Wavelet Transform

A hierarchical planar curve descriptor is developed using the Wavelet Transform, it is mathematical function which divides a time signal into various scale components and it assign frequency range to scale components and it study with resolution and matches that scale [7].

## 2.2 *Color Feature*

Color is one of the important features that make possible to recognize images. Color is generally used to distinguish certain objects, places. Color spaces are RGB, HSV, and HSB [1, 2].

### Color Histogram

This technique is widely used for color features. Color histograms are commonly used as feature vectors for images. It is color distribution of an image. The numbers of pixels' distribution of an image is called histogram. It counts and store a similar pixel. It changes in translation, scaling and easy to compute and insensitive to small variation. It does not take spatial information and it is not robust [8].

### Color Moments

The mean, variance, and standard deviation of an image are called color moment. It compares images based on color and measure similarity between images. It distributes image color by three moments are (1) mean (2) standard deviation (3) skewness. It is rotation invariant and scaling. It is inexpensive to compute [8].

### Color Coherence Vector

Color Coherence Vector is a histogram-based technique for comparing images it is also incorporates spatial information. It divides each histogram into consistent and inconsistent components. It is easy to compute and performance much better than color histograms [9].

### Color Correlogram

CC is not an image partitioning and histogram refinement method. It is used over histogram. Color Histogram divides images into pixels or bins and bins subdivided into 4 bins and these information is stored into the form of correlogram [9].

### Average RGB

Average RGB is an average value of channel RGB. It extracts average values of the channel red, green and blue. It is used to filter out the image. Less computation cost [10].

### Scalable Color Descriptor

It fixes the color space into HSV and quantization of HSV space is 256 bins there is H contains 16 levels and S contains 4 levels, V contains 4 levels [11].

### Color Structure Descriptor

It is used to retrieve image. It represents image by color structure. Using structuring elements, it express color structure and it contains  $8 \times 8$  pixels. It visits all the location of the images [11].

### Dominant Color Descriptor

Dominant color descriptor (DCD) is one of descriptors in MPEG-7, is one of the most commonly used descriptor. It is an effective, compact, and intuitive representation of salient colors in an image region. It consists of n no of dominant colors and it express color component vector and pixel percentage of the image [11].

## 2.3 Edge Features

Edges are set of connected pixels and boundary of regions in image. The edges are horizontal, vertical, and diagonal [11].

### Sobel Operator

The Sobel operator is used to calculate the intensity gradient of the image and to determine the direction of the possible increase from the bright phase to the dark phase and the speed of change of direction. The results show the abrupt or smooth change in intensity of from uses two  $3 \times 3$  cores. Sobel operator is very popular but computational is expensive. We can adjust mask coefficient as per our requirement [11, 12].

$$Mx = \begin{bmatrix} -1 & 0 & -1 \\ -2 & 0 & -2 \\ -1 & 0 & -1 \end{bmatrix} \quad My = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

### Prewitt Operator

The Prewitt operator is similar to the Sobel operator. Prewitt detects horizontal and vertical edges. It calculated the 8 convolution kernels. The convolution kernels with the largest module are selected [13]. This operator uses the kernels:

$$Mx = \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \quad My = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

### Robert Operator

Robert is a simple operator and fast to compute. Because of small mask size it is high sensitive to noise [13]. This operator uses the kernels:

$$Mx = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad My = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

### Laplacian of Gaussian (LoG)

The Laplacian edge detector uses only one kernel. It is used to find out pixel of the dark and light side of the edge. This operator uses the kernels:

$$Mx = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad My = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

It is Gaussian operator uses Laplace operator for the 2nd derivative of the image. Gaussian function used to reduce noise and Laplace operator recognized sharp edge [11, 12].

Gaussian function is

$$G(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\left(\frac{x^2 + y^2}{2\sigma^2}\right)\right) \quad (1)$$

### Canny Operator

It is an optimal edge detector and standard technique. It is derived from the Laplace operator. It is less sensitive to noise. It takes more time and more complex to compute. Localization and detection is good and provide clear response. It uses 2 thresholds for weak edge detection and strong edge detection [11, 12].

## 3 Comparison of Various Image Feature Extraction Methods

The comparison of different image features techniques based on shape, color and edge is discussed in below mentioned tables.

The geometric moments are a very compact shape representation and the computational cost is low, but it is difficult to obtain higher order moment invariants. The Zernike moment is robust to noise. And computational complexity is higher. Fourier descriptors are its robustness, which is able to capture some perceptual features of the shape and are easy to derive. The wavelet transform is available at low cost.

**Table 1** Comparison of different shape methods

Shape methods	Geometric moment	Zernike moment	Fourier descriptor	Wavelet transform
Rotation invariant and translation scale	Yes	Yes	Yes	Yes
Computational complexity	Less	High	Less	Less
Affine transform	Bad	Bad	Bad	Good
Noise resistance, occultation resistance	Noise resistance is bad	Noise resistance is good	Noise resistance is bad	Average occultation resistance

Color histogram is efficient but expensive to compute, color moment is compact and robust and computational cost is less. The CCV contains spatial information and accuracy is better than histogram. The color correlogram captures the spatial correlation only between identical colors. Average RGB has less computation than others. The performance of CSD depends on structure and size of the windows and it is difficult to determine. SCD needs post processing for spatial information. DCD is more valid and compact (Table 3).

The performance of the Canny operator better than other operator. The Canny algorithm is more computationally expensive than the Sobel, Prewitt and Robert operator. Sobel also detects various features and is more computationally efficient as Canny, but with more false edges. Robert operator is very sensitive to noise but faster to compute. Laplacian is easy to edge detection and orientation. Prewitt is best for orientation.

## 4 Conclusion

As the comparison of different techniques are studies from various standard published survey and research papers, the study varies based on the results and parameters. The motivation is to obtain the best technique to be used in further data extraction from images. The process of image processing is performed by keeping the extracted information and required data in consideration. the black and white image processing are processed with a wavelet based transformation process and RGB images are processed with kernel based transformation which can generate different color channels based on histogram. The main aim of this survey was to clarify the purpose of each of the popular techniques for color, feature and shape extraction of images irrespective of its type and size.

**Table 2** Comparison of different color methods

Color method	Color histogram	Color moments	Color coherence vector	Color correlogram	Average RGB	Scalable color descriptor (Scd)	Color structure descriptor (Csd)	Dominant color descriptor (Dcd)
Spatial information	No	No	Yes	Yes	No	Need post processing for spatial information	Yes	No
Noise	Sensitive to noise	Less	Less	Sensitive to noise	Less	Less	Sensitive to noise	Less
Computation cost	Expensive to compute	Less	High	Very high	Less	Less	Complex computation	Inexpensive
Accuracy	Better	Consistently out performs	Average, better than color histogram	Good	Less accurate if not combined with other feature	Good	Better	Accurate if compact

**Table 3** Comparison of different edge methods

Edge methods	Sobel	Prewitt	Laplacian	Robert	Canny
Computation	Simple and time efficient	Less expensive, computations like Sobel	Faster to calculate	Very quick	Complex and time consuming
Noise	Less sensitivity	Least sensitivity	Least sensitivity	Very sensitive	Less sensitive
Detection of edges and orientation	Very easy to search smooth edge	Best operator detects the orientation	Easy	Very easy	Robust approach for edge detection
Diagonal direction points	Not preserved always	Not preserved always	N/A	Preserved	N/A
Performance	Does not give appropriate result. Detect with rough and thick edge	Good for detecting vertical and horizontal edge, faster method for edge detection	Fixed characteristics in all directions	Performs a simple, 2-d spatial gradient measurement on an image, highlight changes in intensity in a diagonal direction	Good performance of detecting edge, extract image features without altering the features
Accuracy	Not very	Not very	High above average	Not very	Accurate
Time complexity	Low	Low	Low	High	High

## References

1. A.O. Salau, S. Jain, Feature extraction: a survey of the types, techniques, applications, in *International Conference on Signal Processing and Communication (ICSC)* (2019)
2. J.Y. Huang, C.S. Wang, T.K. Shih, An intelligent content based image retrieval system based on color, shape and spatial relations, in *Proceedings of the National Science Council, R. O.C. Physical Science and Engineering* (2001)
3. J. Liu, Y. Liu, Image feature extraction method based on shape characteristics and its application in medical image analysis. *Appl. Inf. Commun.* 172–178 (2011)
4. Y. He, B. Ye, M. Liu, Image Zernike moments shape feature evaluation. *Geo-spatial Inf. Sci.* **10** (3), 191–195 (2007)
5. R.J. Prokop, A.P. Reeves, A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graph. Models Image Process* **54** (5), 438–460 (1992)
6. A.P. Reeves, R.J. Prokop, S.E. Andrews, F.P. Kuhl, Three-dimensional shape analysis using moments and Fourier descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **10** (6), 937–943 (1988)

7. Y. Mingqiang, K. Kidiyo, R. Joseph, A survey of shape feature extraction techniques. *Pattern Recogn. Tech. Technol. Appl.* (2008)
8. A.K. Jain, A. Vailaya, Image retrieval using color and shape. *Pattern Recogn.* **29** (8), 1233–1244 (1996)
9. F. Long, H. Zhang, D.D. Feng, Fundamentals of content-based image retrieval. *Signals Commun. Technol.* 1–26 (2003)
10. A. Chadha, S. Mallik, R. Johar, Comparative study and optimization of feature-extraction techniques for content based image retrieval. *Int. J. Comput. Appl.* **52** (20), 35–42 (2012)
11. B.S. Manjunath, J.R. Ohm, V.V. Vasudevan, A. Yamada, Color and texture descriptors, in *IEEE Transactions on Circuits and Systems for Video Technology* (2001), pp. 703–715
12. C. Harris, M. Stephens, A combined corner and edge detector, in *Fourth Alvey Vision Conference*, Manchester, UK (1988)
13. R. Muthukrishnan, M. Radha, Edge detection techniques for image segmentation, *IJCSIT* (2011)
14. M. Kumar, R. Saxena, Algorithm and technique on various edge detection: A survey, *SIPIJ* (2013), pp. 65–75
15. H. Lee, S. Jeon, I. Yoon, J. Paik, Recent advances in feature detectors and descriptors: A survey **5** (3) (2016)

# Cyber Security Threat, it's Risks and Susceptibilities, in the Global Perspective



Radhika Mohan Gupta, Priyanka Mathur, and Ruchi Nanda

**Abstract** Cyber threats have become a big problem for the world today and out of many, the most prominent in this respect are social engineering and cyber warfare. Considering cyber security, social engineering accomplishes its malicious goals by captivating human weaknesses. Social engineering is a grim security threat to nations, users, data and its operations and technological infrastructure. It has numerous conceptual flaws (for instance conflicting conceptual intentions, nebulous conceptual boundaries, confusing instances, abuse etc.) thus creating a negative impact on the holistic understanding of social engineering attacks. Cyber warfare is the new threat the nation state system is facing, and social engineering is one of the ways to achieve that. In this paper, a systematic literature review is conducted and the conceptual progression and technical development as a result of social engineering and cyber warfare is analysed systematically, along with measures to evade such attacks. User awareness as a solution to this camouflaged attack is also suggested. This research, then, aims to address these conceptual problems and propose a new and fresh understanding of social engineering and cyber warfare in cybersecurity (SEiCS). The literature review leads to conclude that cyber security not only needs to be monitored from technological perspective, but a more holistic approach is required to gauge its defence.

**Keywords** Social engineering · SEiCS · International affairs · Cyber security awareness

---

R. M. Gupta · P. Mathur (✉)  
Manipal University, Jaipur, India  
e-mail: [priyanka.mathur@jaipur.manipal.edu](mailto:priyanka.mathur@jaipur.manipal.edu)

R. M. Gupta  
e-mail: [radhikamohan.gupta@jaipur.manipal.edu](mailto:radhikamohan.gupta@jaipur.manipal.edu)

R. Nanda  
IIS Deemed to be University, Jaipur, India  
e-mail: [ruchi.nanda@iisuniv.ac.in](mailto:ruchi.nanda@iisuniv.ac.in)

## 1 Introduction

The world has probably not seen a technological revolution like that of the internet revolution where the virtual world is becoming a strong alternative for nearly everything that is taking place in normal life. This, can then, also be seen in the case of warfare. Everyone is familiar with the traditional use of warfare that includes the physical engagement of militaries in some form or another. Cyber warfare refers to taking this conflict into the virtual world where the internet is used to attack. Thus, the internet, which is being put to much good use in today's world, is also being used to cause chaos using cyberspace.

The world has always faced traditional threats emanating from military and border insecurities. This has been the primary reason why the world has seen wars of the scale of the world wars. However, in the changed world scenario of today, various other threats are emerging that are disturbing the security of nation states in much the same manner as military threats do. These are called nontraditional security threats. Some important non-traditional security threats that have cropped up in recent times are food and water crises, energy insecurities, environmental problems, terrorism and the like. "Among these, cyber threat is the most advanced, complicated and ungoverned in the realm of interstate relations in view of technological advancement, lawlessness of cyberspace and widening scope of the arena. Thanks to globalization, advances in information and communication technologies (ICTs) are diffusing across societies and internet is becoming a universal and prominent phenomenon" [1].

A cyber-attack can affect technological infrastructure through various ways like through viruses, malware, or can affect personnel through social engineering or cyber bullying [2]. Mouton states that "Social engineering is the art of manipulating human weaknesses in order to attain a malicious goal" [3]. Social engineering attacks are on a rise exponentially and individuals, corporate houses and organizations are falling prey to it and their classified information and/or valuable assets are at risk by manipulating legitimate users. The attackers exploit human vulnerabilities which are very difficult to secure [4, 5].

In cyber security issues, social engineering paves the way for severe and tenacious security threats. By the end of the year 2020, a novel yet simple scam was identified that manipulated Google Drive's notification system. In this scam a document was created which had malicious links to few phishing sites. The scammer would target the victim by commenting in the document and tagging the victim. Then the victim would be asked to collaborate. After being tagged the victim gets an email notification from Google which has the comment's text and a redirection link to the document. Falling in the scam the target views and reads the document and feels elated on being asked to collaborate. Resultantly they would click on the malicious links and get redirected to the phishing site. They would then enter their login credentials or other personal data, making them vulnerable. This scam was tricky as it had exploited Google's email notification system for earning legitimacy [5]. Such scams pose a severe threat to online personal identity. This is how social engineering manipulates the humane aspect of technology through deceit [6, 7]. It

uses communication between humans as a medium to enter in a system in forbidden manner [5]. Such attacks can happen through psychological i.e., human-based attack as well as through physical level i.e., computer based attack [8, 9]. People will be convinced to disclose their sensitive information or may even be asked to provide access to their personal computer systems. Eventually through social engineering digital as well as physical security is evaded, and human manipulation is being done. Along with it another threat is evasion in cyberspace. Hackers or adversaries find it difficult to attack networked and independent computer systems and networks [10]. Consequently, hackers count on social engineering attacks to circumvent technical controls. To attack countries, organizations and even individuals these attacks are conducted either on one-to-one basis which is unequivocal or over phone or even via computer-interaction through malicious websites, instant messages, pop-up windows or emails.

Here, psychological manipulation of the users is the modus operandi in use and unauthorized access to their system is gained as a result [11, 12]. There have been measures taken to increase the level of employee awareness against these attacks and Mouton [3] highlighted the importance for the same.

## 2 Cyberspace and Its Impact Upon International Affairs

The advancement in information and communication technology since the 1980s, the birth of the internet in that decade, the worldwide spread of the internet at a furious pace, thereafter and the adoption of the internet by people in their day to day lives to the extent that internet is now a critical part of the lives of human beings has all created a completely new environment for the world to live in: cyberspace. “Physically, cyberspace can be defined as the digital medium’s extension across various lines of transmission—metallic, fibre optic, and wireless—and their channels on the network of networks, the Internet. In these terms, cyberspace is the technological expression of the information super-highway” [13].

While the internet has solved many a problem in today’s world and is brilliantly useful to individuals and governments alike, it is being used to create unrest within nation states also which has the potential to turn into worldwide conflicts. While traditional issues like military build-up have caused security dilemma between nation states in the past, nations today are facing cyber insecurity wherein cyber development in nations is being watched by other nations as this now has the potential of becoming a state threat. “The most common cyber activity among the states is espionage for political, security and economic purposes. In the military arena, cyberspace would potentially prove a game changer as it is redefining the traditional concepts of strategic stability, strategic forces, deterrence, etc., not least among major powers” [1].

The nature of cyber threats “range from non-kinetic such as cybercrime, espionage, hacking and information propaganda to kinetic viz. cyber warfare” [1]. According to Joseph S. Nye, “cyber war and economic espionage are largely associated with states and cyber crime and cyber terrorism are mostly associated with

nonstate actors” [14]. Moreover, attacks can be from multiple agencies: state and non-state actors alike. States can launch cyber warfare against other states with whom they have inimical relations in order to weaken them internally. However, these can be carried out by almost anyone, individual, organization or a state. There is always a great threat of this medium being used by terrorist and other anti-state bodies to destabilise governments, to spread terror and propaganda and cause general distress of varying intensities.

According to Timothy Shimmell, the intensity of a disaster caused by a cyber-attack are threefold. He says that there are “several levels of cyber war, of which three stand out: cyber war as an adjunct to military operations; limited cyber war; and unrestricted cyber war” [15]. The first, says Shimmell, aims at incapacitating the cyber capabilities of enemy armies in order to weaken their military defences. The second kind of cyber war has a limited scope but causes damage. The third kind, i.e. the unrestricted cyberattack is the deadliest because it causes intended physical destruction at a very big scale, loss of life to property, destabilizes infrastructure and damages economic and social life [15].

The probability of such attacks on nation states are seen to be high, owing to the vulnerable position of these nation states because governments use the cyberspace in various critical domains, e.g. political, military, economic social or any other referring to the all-encompassing nature of the cyber space and the national security of a country and its people. The 2013 White Paper of the Australian government has called cyberattacks a threat to defense, government and commercial information networks [16].

The cyber domain is the new age war zone now, apart from the traditional domains of land, sea, air and outer space [17]. Nation states see cyber threat as a nontraditional form of threat. They are building mechanisms to prevent against the same. A clearer perspective of the same has allowed states like Australia to increase “its intrusion detection, analytic and threat assessment capabilities and improved its capacity to response to cyber security incidents” [18].

There is a need for mutual assistance between nation states for responding to cyber crises “but the effectiveness of such cooperation depends greatly upon strategically aligned policy goals and bilateral and multilateral relations” [19].

Cyber-attack is seen as a global threat by the United Nations [17]. For this reason, there is a build-up in multilateral dialogue happening in the world on cybersecurity issues.

“For example, the United Nations First Committee has been actively examining the Developments in the Field of Information and Telecommunications in the Context of International Security for years. The African Union has published the Draft African Union Convention on the Establishment of a Credible Legal Framework for Cyber Security in Africa. The European Union (EU) has recently published a Joint Communication on the Cyber Security Strategy of the European Union, which is the first attempt for a comprehensive EU policy document in this domain to reflect the common view on cybersecurity of all its 27 member states” [19].

### 3 Review Methodology

A systematic literature review (SLR) is conducted in this research which focuses on social engineering attacks and cyber warfare across nations, challenges in detecting these attacks and the defence mechanism that can safeguard stakeholders against these types of attacks.

A comprehensive systematic literature review in the research field of social engineering and cyber warfare and their global spread is done and shown in Table 1, the objective of this research is to synthesize existing knowledge. The methodology of this research consists of two stages. The first phase includes searching for relevant literature in both fields: social engineering and cyber warfare and their global spread and relevance today. The other phase is to analyze the selected literature to identify the most common social engineering threats and the ways to evade them.

**Table 1** Research questions and approaches

Research question	Research direction
Who are more vulnerable to social engineering attacks?	Subconscious mind [20, 21], employees with BYOD (bring your own device) [22–24], new employees [3, 25, 22, 26], Clients [25, 26], Partners and Contractors [25, 26], Top level Management [25, 26], IT Professionals [25, 26]
What are different social engineering attack strategies?	(a) The Physical Locations: Workplace [20], Telephone [27], Online [28] (b) The Psychological Methods: Authority to gain vital information [3], exploiting natural inclination to help [21], developing undue personal connections, showing unwarranted commitment and consistency, forceful reciprocation [28], targeting low Involvement people [29]
How to educate the users against such attacks?	Security Policy [30], Education and Training [25], Network Guidance [31], Audits and Compliance [32], Technical Procedures [32], Physical Guidance [32]
Which is the best method amongst all alternatives?	Education and Training [3, 25, 32]
Cyber warfare as a new tool of warfare	A non-traditional threat to the international system operating with multiple methods [1], Cyber terrorism by non-state actors [14] Has tremendous destructive effects [15]
Cyber threat as a global threat to which the nation-state international system is beginning to respond	Collective recognition of threat, international cooperation to contend the threat [17, 19]

## 4 Conclusion

The world is under attack, by a new non-traditional threat, social engineering along with cyber warfare. Cyber warfare is becoming the new danger to the security of nation states as it does not require traditional methods of warfare and can be carried out by the click of a button from anywhere and can still have an explosive impact if intended. Nation states are now seen to be growing aware of this new threat to their security and the international system is recording a flurry of deliberations in the recent past and present that are trying to find ways and means to battle this menace. Cyber warfare plays out in multiple ways and one of the most camouflaged and yet very dangerous is the means of social engineering. Social engineering has the capacity to affect individuals, thereby affecting nations and the nation state system as a whole. Effective ways to evade such attacks is primarily user awareness. All new employees of an organization need to be made aware of such attacks. Although technology tries to reduce this, yet vulnerability is an integral component of human behavior. These psychological predispositions can be influenced only through education. Organizational education is suggested but an effective and absolute solution to these threats is yet to be put forward.

## References

1. M.R. Shad, Cyber threat in interstate relations: case of US-Russia cyber tensions. *Policy Perspect.* **15**(2), 41–55 (2018)
2. D. Saratchandra, K. Haltinner, N. Lichtenberg, College students' cybersecurity risk perceptions, awareness, and practices, in *2016 Cybersecurity Symposium (CYBERSEC)* (IEEE, 2016), pp. 68–73
3. F. Mouton, L. Leenen, H.S. Venter, Social engineering attack examples, templates and scenarios. *Comput. Secur.* **59**, 186–209 (2016)
4. D. Kahneman, *Thinking, Fast and Slow* (Penguin Books, 2011)
5. Z.L. Švehla, I. Sedinić, L. Pauk, Going white hat: Security check by hacking employees using social engineering techniques, in *2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (IEEE, 2016), pp. 1419–1422
6. S. Uebelacker, S. Quiel, The social engineering personality framework, in *2014 Workshop on Socio-Technical Aspects in Security and Trust* (IEEE, 2014), pp. 24–30
7. A.S. Alazri, The awareness of social engineering in information revolution: techniques and challenges, in *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)* (IEEE, pp. 198–201) (2015)
8. G.L. Orgill, G.W. Romney, M.G. Bailey, P.M. Orgill, The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems, in *Proceedings of the 5th Conference on Information Technology Education* (2004), pp. 177–181
9. I. Ghafir, V. Prenosil, A. Alhejailan, M. Hammoudeh, Social engineering attack strategies and defence approaches, in *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)* (IEEE, 2016), pp. 145–149
10. S.D. Applegate, Social engineering: hacking the wetware! *Inf. Secur. J. Glob. Perspect.* **18**(1), 40–46 (2009)
11. C. Hadnagy, *Social Engineering: The Art of Human Hacking* (Wiley, 2010)

12. A. Berg, Cracking a social engineer. *Comput. Secur.* **8**(14), 700 (1995)
13. K.W. Mahmoud, *Cyber Attacks: The Electronic Battlefield* (Arab Center for Research and Policy Studies, 2013). <http://www.jstor.org/stable/resrep12651>. Last accessed on 21 Nov 2021
14. J.S. Nye, Nuclear lessons for cyber security? *Strateg. Stud. Quart.* **5**(4), 18–38 (2011)
15. S. Timothy et al., Countering cyber war, *NATO Rev.* (2001). <https://www.nato.int/docu/review/articles/2001/12/01/countering-cyber-war/index.html>. Last accessed on 25 Oct 2021
16. D. Ball, G. Waters, Cyber defence and warfare. *Secur. Chall.* **9**(2), 91–98 (2013)
17. N. Melzer, *Cyberwarfare and International Law* (UNIDIR Resources, 2011). <https://unidir.org/files/publications/pdfs/cyberwarfare-and-international-law-382.pdf>. Last accessed on 22 Nov 2021
18. Commonwealth of Australia, *Defence White Paper 2013* (Canberra Department of Defence 2013)
19. A.-M. Taliharm, *Towards Cyberspace: Managing Cyber War Through International Collaboration*. <https://www.un.org/en/chronicle/article/towards-cyberpeace-managing-cyberwarthrough-international-cooperation>. Last accessed on 26 Nov 2021
20. I. Mann, *Hacking the Human: Social Engineering Techniques and Security Countermeasures* (Routledge, 2017)
21. J. Goodchild, Social engineering: the basics. *CSO Online* (2012)
22. K. Krombholz, H. Hobel, M. Huber, E. Weippl, Advanced social engineering attacks. *J. Inf. Secur. Appl.* **22**, 113–122 (2015)
23. M.M. Singh, S.S. Siang, O.Y. San, N. Hashimah, A.H. Malim, A.R.M. Shariff, Security attacks taxonomy on bring your own devices (BYOD) model. *Int. J. Mob. Netw. Commun. Telemat. (IJMNCT)* **4**, 117 (2014)
24. A.G. Bello, D. Murray, J. Armarego, A systematic approach to investigating how information security and privacy can be achieved in BYOD environments. *Inf. Comput. Secur.* (2017)
25. M.R. Arabia-Obodoza, G. Rodriguez, A. Johnston, F. Salahdine, N. Kaabouch, Social engineering attacks a reconnaissance synthesis analysis, in *2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)* (IEEE, 2020), pp. 0843–0848
26. F. Salahdine, N. Kaabouch, Social engineering attacks: a survey. *Future Internet* **11**(4), 89 (2019)
27. J.W. Bullée, L. Montoya, M. Junger, P.H. Hartel, Telephone-based social engineering attacks: an experiment testing the success and time decay of an intervention, in *Proceedings of the Singapore Cyber-Security Conference (SG-CRC) 2016* (IOS Press, 2016), pp. 107–114
28. D. Irani, M. Balduzzi, D. Balzarotti, E. Kirda, C. Pu, Reverse social engineering attacks in online social networks, in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment* (Springer, Berlin, Heidelberg, 2011), pp. 55–74
29. D. Kvedar, M. Nettis, S.P. Fulton, The use of formal social engineering techniques to identify weaknesses during a computer vulnerability competition. *J. Comput. Sci. Coll.* **26**(2), 80–87 (2010)
30. H. Aldawood, G. Skinner, Contemporary cyber security social engineering solutions, measures, policies, tools and applications: a critical appraisal. *Int. J. Secur. (IJS)* **10**(1), 1 (2019)
31. E. Blancaflor, C.V.H. Banzon, C.J.J. Jackson, J.N. Jamena, J. Miraflores, L.K. Samala, Risk assessments of social engineering attacks and set controls in an online education environment, in *2021 3rd International Conference on Modern Educational Technology* (2021), pp. 69–74
32. N.Y. Conneh, P.J. Schmick, Cybersecurity: risks, vulnerabilities and countermeasures to prevent social engineering attacks. *Int. J. Adv. Comput. Res.* **6**(23), 31 (2016)

# Detecting Cheating in Online Assessments Using Face Recognition and Head Pose Estimation



Shaik Mohissin Sultana, M. Kameshwara Rao, and S. G. Santhi

**Abstract** MOOCs (massive open online courses) and other forms of distant learning are becoming more popular and widespread. In this next stage of education's scalability, the ability to efficiently proctor remote formative assessments is a critical limiting factor. Human proctoring, which entails either compelling test takers to visit an examination center or watching them visually and vocally throughout tests through a webcam, is now the most common technique of assessment. Such techniques, on the other hand, are time consuming and costly. A camera and microphone are all part of the system's hardware.

**Keywords** Online proctoring · Face detection · Face recognition · Head pose detection

---

S. M. Sultana (✉) · M. K. Rao

Department of Electronics and Computer Engineering, KLEF, Vaddeswaram, India  
e-mail: [180050043@kluniversity.in](mailto:180050043@kluniversity.in)

M. K. Rao

e-mail: [kamesh.manchiraju@kluniversity.in](mailto:kamesh.manchiraju@kluniversity.in)

S. G. Santhi

Department of Computer Science Engineering, Annamalai University, Chidambaram, India

## 1 Introduction

A variety of educational courses and training are available at many institutions and universities in physical classrooms and places, with lectures, admissions examinations, end exams, and other operations requiring attendance. Featuring the ability for people to take lessons from the comfort of their own homes, at their own time and location, reduced costs, a far lower environmental impact, and many more. Indeed, online learning has the ability to alleviate or minimize all of the above-mentioned issues with in-class training. This paper describes a ground-breaking online monitoring system that employs deep learning to continually monitor physical locations without requiring the presence of a physical monitor. Among the biometric approaches employed by the system are facial recognition using the HOG face detector and the OpenCV face recognition algorithm. The Face Detection Data Set and Benchmark (FDDB) and LFW (Labeled Faces in the Wild) datasets are used to test the system, which is implemented as a software system.

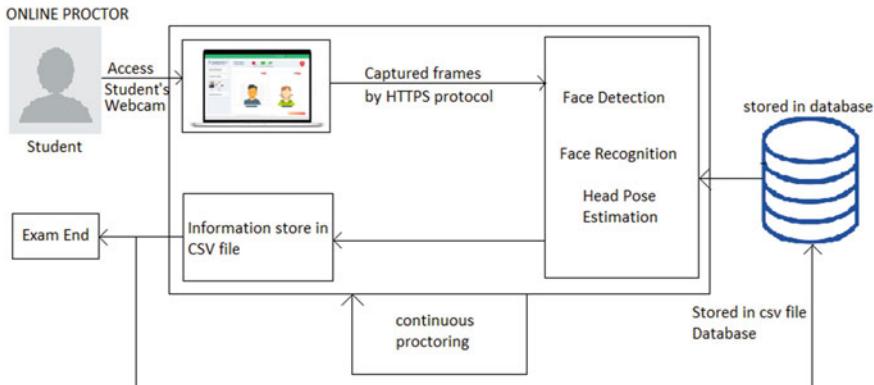
## 2 Related Works

Throughout the test, the online exam faces several hurdles. Discussed the online exam's many issues and proposed a solution that included grouping client hostnames or IPs for a given region and time, facial recognition and fingerprints, as well as a biometric solution. A profile-based authentication system architecture for the on-screen test is provided, which is based on many demanding questions, such as a personal question, a favorite question, and a professional question. In 2018, a multi-biometric continuous authentication system that recognized the user's face, voice, touch, mouse, and keyboard. For intrusion detection, builds and installs a network firewall system that uses Network Address Translation, Demilitarized Zone, and Virtual Protocol Network as well as other firewall technologies. A deep learning-based face identification also a solution for detecting the protected online test. A introduction of another An online exam management system that uses fingerprint and eye tracking technology. Students' cheating and non-cheating status are utilized to assess their recommended methods. For researchers and engineers, the online proctoring system has recently become a difficulty. The demand for online monitoring is steadily increasing, as are the obstacles as a result of the Coronavirus issue. Several industrial enterprises commercialized the proctoring system with the paid version. This is the entire process of Online Proctoring, from the beginning of the test to the finish of the exam. Artificial intelligence (AI) have changed our environment and surroundings by incorporating intelligence into many of our daily chores [1–3], yet it has several downsides [4, 5]. Education [6], transportation [7], communication networks [8], disaster management [9], smart cities [10], and many more fields have benefited from machine and deep learning innovation. AI has the ability to completely transform Proctoring and online learning.

### 3 Proposed Model

There are two parts to the proposed web-based online proctoring system. First, there's the online proctoring. The suggested design for the online proctoring system is shown in Fig. 1.

During the exam, the online proctoring software verifies the examinee's identity and prevents any suspicious conduct. The program guarantees that the participant shares a screen with video and audio recording even before the test begins. The test does not begin until the proctors have verified their identities. The major topic begins with online proctoring, in which a student takes an exam from a monitor using the camera of the student. The webcam uses the HTTPS Protocol to capture frames and detects three types of techniques: Face Detection, which detects the face using facial landmarks, Face Recognition, which determines whether the face matches the student or not using loops, and Head Pose Estimation, which estimates the student's head movement and angle using facial landmarks and some functions and algorithms. When the data is discovered, it will be saved in a CSV file. This method will be followed throughout the exam. The test will be proctored continuously during this procedure. The data will be saved in a database as a CSV file. This procedure will conclude once the entire inspection has been completed, and the collected data will be saved in a CSV file database.



**Fig. 1** Proposed architecture

**Algorithm 1 Master Algorithm**

---

```
1: procedure online Proctor
2:   while True do
3:     frame ← captureFrameFrom WebCam
4:     getfaceDetect ← getfaceDetection(method "HOG")
5:     findface ← findFacetForCurrentFrame
6:     facelandmarks ← findfacelandmarks for current frame
7:     detectmarks ← findfacemarks
8:     if getface == 0 then
9:       cancel Exam: face not found
10:    else if getface == 1 then
11:      continue Exam
12:    end if
13:    face Match ← faceRecognition()
14:    if face Match == False then
15:      cancel Exam: Unjustified face
16:    end if
17:    angle ← HeadPoseDetection()
18:
19:  end if
20:
21: end while
22:
23: end procedure
```

---

***Face Recognition***

The most extensively used biometric for internet authentication is face recognition. Intel created the OpenCV computer vision library in 1999. Image representation, Affect its operations, and LBPH (Local Binary Patterns Histograms) are some of the face recognition techniques supported by OpenCV. These methods are based on the Cascade classifier Haar technology developed by Paul and Michael. We utilise a photograph of tutee as the input and employ HOG methods to detect facial landmarks in the picture in our suggested methodology. Estimate the 68 landmarks for the identified picture. The photos are collate to faces we've already identified and saved in our database. Algorithm 2 shows the facial recognition pseudo code.

**Algorithm 2 Recognition of Face**

---

```
1: procedure Face Identification (Id, Name)
2:   while True do
3:
4:     Case ← present Case
5:     Locateface ← get all faces
6:     Encodingface ← get all faceEncoding
7:     faceMatch ← compare all faces with student face
8:
9:     if sameface == True then
10:
11:       Face same
12:
13:     else
14:
15:       Face not same
16:
17:     end if
18:
19:   end while
20:
21: end procedure
```

---

***Head Pose Detection***

The six areas of the face that we require using traditional 3D coordinate of face characteristics, we estimated the rational and translational vectors at the nose tip. We now need to know the camera's intrinsic properties, such as focal length, optical centre, and radial distortion parameters, in order to make an accurate estimate. To make our task easier, we may estimate the first two and pretend the last one isn't present. After we have the necessary vectors, we can project those 3D points onto a 2D surface, which is our picture. This is known as the Perspective-n-Point problem, or PNP, in computer vision jargon. The goal of this job is to determine an item's posture. Translation is process of moving the camera from its present 3D location ( $X, Y, Z$ ) to a new 3D place ( $X', Y', Z'$ ). Translation has 3 degree freedom. We can go in one of three directions: X, Y, or Z. The vector  $t$  represents translation and is equal to  $(X' - X, Y' - Y, Z' - Z)$ . The camera may also be rotated around the X, Y, and Z axes. To estimate a 3D object's posture, you'll need to discover six numbers: 3 to translation and 3 to rotation.

---

**Algorithm 3 Head Pose Detection Algorithm**


---

```

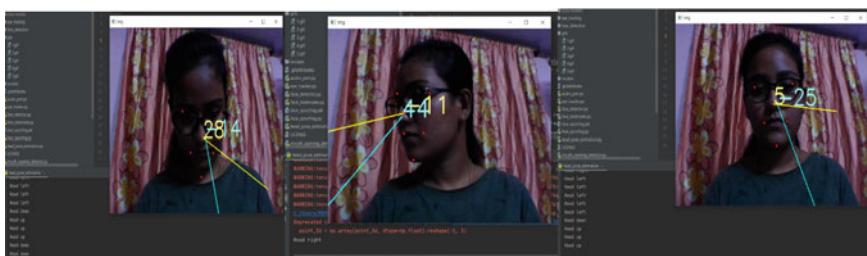
1: procedure get_2d_points(rotationvector,translationvector,cameramatrix,)
2: while True do
3:   frame ← currentFrame
4:   getheadposepoints ← frame.img
5:   facemodel ← getfacedetector()
6:   landmark_model ← get_landmark_model()
7: while True
8:   If ret == True
9:     faces ← find_faces(img,face_model)
10:    for p ← image_points
11:      if ang1>=48:
12:        Print('Head Down')
13:      elif ang1<= -48:
14:        Print('Head Up')
15:      If ang2 >=48:
16:        Print('Head Right')
17:      elif ang2 <= -48:
18:        print('Head Left')
19:    cv2.waitKey(1)
20:    end if
21:  end for
22: end while
23: end procedure

```

---

## 4 Results

See Fig. 2.



**Fig. 2** Cheating Detection using head pose

## 5 Conclusion

In order to provide complete information for online assessments, this paper uses face recognition and head pose detection algorithms. During the online exam, our proposed strategy will assist in decreasing disparity. When using an online proctoring system, human-induced detection is critical since it helps spot suspicious activity among students throughout the exam. Because we only have Head pose Detection and face recognition algorithms for online monitoring, further, we would like to employ and analyze more human behaviors like Multi Person Detection, conversing with others, focusing on different angles and so on. In the bulk of the tests, with a high number of false prediction positives. Dlib and MTCNN has extremely comparable outcomes, with MTCNN having a minor advantage, while Dlib is unable to recognise very tiny faces. Also, if the picture dimension is large, with minimal closure. MTCNN may produce the good output as Fig. 2, as we saw while contrasting photos. It can distinguish side faces and works well with occlusion and fast head motions. It also had the fastest frame rate. An administrator can administer the exam from a remote place with the aid of online remote proctoring. Proctoring guarantees that online assessments are cheat-proof, allowing any applicant taking the test from their own home to complete it without cheating. Two datasets were used to test our suggested system. Further examine and evolution of the system in a real-world circumstance will be the focus of future work. This paper aims to create unlock structure so that the section may educate from one another, resulting in faster breakthroughs in the area as a result of open-source advances.

## References

1. T. Yigitcanlar, N. Kankanamge, M. Regona, A. Maldonado, B. Rowan, A. Ryu, K.C. Desouza, J.M. Corchado, R. Mehmood, R.Y.M. Li, Artificial intelligence technologies and related urban planning and development concepts: how are they perceived and utilized in Australia? *J. Open Innov. Technol. Mark. Complex.* **6**(4), 187 (2020) (Online). Available: <https://www.mdpi.com/2199-8531/6/4/187>
2. T. Yigitcanlar, L. Butler, E. Windle, K.C. Desouza, R. Mehmood, J.M. Corchado, Can building “artificially intelligent cities” safeguard humanity from natural disasters, pandemics, and other catastrophes? an urban scholar’s perspective. *Sensors* **20**(10), 2988 (2020) (Online). Available: <https://www.mdpi.com/1424-8220/20/10/2988>
3. N. Janbi, I. Katib, A. Albeshri, R. Mehmood, Distributed artificial intelligence-as-a-service (DAIaaS) for smarter IoE and 6G environments. *Sensors* **20**(20), 5796 (2020) (Online). Available: <https://www.mdpi.com/1424-8220/20/20/5796>
4. T. Yigitcanlar, J.M. Corchado, R. Mehmood, R.Y.M. Li, K. Mossberger, K. Desouza, Responsible urban innovation with local government artificial intelligence (AI): a conceptual framework and research agenda. *J. Open Innov. Technol. Mark. Complex.* **7**(1), 71 (2021) (Online). Available: <https://www.mdpi.com/2199-8531/7/1/71>
5. T. Yigitcanlar, R. Mehmood, J.M. Corchado, Green artificial intelligence: towards an efficient, sustainable and equitable technology for smart cities and futures. *Sustainability* **13**(16), 8952 (2021) (Online). Available: <https://www.mdpi.com/2071-1050/13/16/8952>

6. F. Alam, A. Almaghthawi, I. Katib, A. Albeshri, R. Mehmood, iResponse: an AI and IoT-enabled framework for autonomous COVID-19 pandemic management. *Sustainability* **13**(7), 3797 (2021) (Online). Available: <https://www.mdpi.com/2071-1050/13/7/3797>
7. M. Aqib, R. Mehmood, A. Alzahrani, I. Katib, A. Albeshri, S. Altowaijri, Smarter traffic prediction using big data, in-memory computing, deep learning and GPUs. *Sensors (Switzerland)* **19**(9) (2019)
8. E. Alomari, I. Katib, A. Albeshri, R. Mehmood, COVID-19: detecting government pandemic measures and public concerns from twitter arabic data using distributed machine learning. *Int. J. Environ. Res. Publ. Health* **18**(1), 282 (2021) (Online). Available: <https://www.mdpi.com/1660-4601/18/1/282>
9. T. Muhammed, R. Mehmood, A. Albeshri, I. Katib, UbeHealth: a personalized ubiquitous cloud and edge-enabled networked health-care system for smart cities. *IEEE Access* **6**, 32258–32285 (2018)
10. R. Mehmood, F. Alam, N.N. Albogami, I. Katib, A. Albeshri, S.M. Altowaijri, UTiLearn: a personalised ubiquitous teaching and learning system for smart societies. *IEEE Access* **5**, 2615–2635 (2017)

# Opinion System of Real-Time Emotions from Human Interaction Incorporated in Class Using Deep Learning



M. Bindhu, S. Asha, and M. Praveen

**Abstract** The paper represents sentiment of recognition or identification of human facial expression used in enormous application like feedback system security, medical diagnosis. The colleges usually uses manual feedback system collected at every semester to know the student/audience response. The collected data are processed manually and consolidated evaluation sheet is prepared with student response to each subject and staff. Hence to provide a automated real time feedback method a deep learning methodology is analysed with Convolution Neural Network (CNN).This algorithm helps in capturing, analyzing and detect emotions from facial expressions of each student. The emotions are classified into four categories as happy, sad, anger and neutral. The classifier use Haar in Keras high level modelling system where the network runs on top of tensor flow framework. The collective data and results obtained using this opinion poll real time system clearly impress the effective feedback system under human emotions.

**Keywords** Convolution neural network · Deep learning · Emotion classification

## 1 Introduction

The emerging and popular technology in the present-day world is Machine Learning. Image processing is one such application where machine learning algorithms are applied in Medical diagnosis, Space technology, Data centres. Facial expression is a non-verbal gesture which signifies the emotional state of a human apart from hand-crafted features [1]. Detection of Emotion from facial features is used in artificial intelligence, machine vision and robotics applications. In this Deep Learning based FER overheads and reliance on the face-based models uses convolutional layers sideways learners direct from the input images. The most admired networks is Convolutional Neural network (CNN), where a weight from each layer of input mages are convolved with filter collection. The intention of this paper is to extend Classroom

---

M. Bindhu (✉) · S. Asha · M. Praveen  
SEC, Chennai, India  
e-mail: [deepbind@gmail.com](mailto:deepbind@gmail.com)

Feedback system using Facial emotions and reactions which can be taken from the human facial expressions as samples input, distinguish and catalogue it to different expression class to map the feedback individual student.

## 2 Literature Review

Shojaeilangari et al. [2] proposes an approach extreme sparse learning, joint data set of root nonlinear classification model. Results were also analysed for unwanted noise and non relation data present in background.

Neethu et al. [3] paper discuss on sentiment analysis with polarity nature. Uses Random forest technique with exactness 90%.

Wati et al. [4] discuss on the performance of face detection features are of human color, light angles, distance and power. The face detection works on distance person and camera for 240 cm.

Zhao et al. [5] highlights on reviews of speech emotion and potential future issues in speech recognition.

Bhatti et al. [6] the paper represents experiments conducted on a presently face expression measured from instructor uploaded as dataset along with triplet benchmark dataset by the Japanese Female Facial Expression (JAFFE) dataset with RELM of 5 teachers accuracy.

Dzedzickis et al. [7] the paper covers IOT system on sensors of contactless methods, contact method which uses penetrating electrodes to observe emotion and intensity levels.

Kuruvayil and Palaniswamy [8] the Emotion Recognition using Meta-learning across Occlusion, Pose and Illumination method acts on human affect identification with multi level PIE data base learning with 68% accuracy.

Hakim [9], the proposed NFC used here identify information/command from a tag where data is observed in near proximity or by touching sense.

Naik [10], this paper is of wireless attendance management system using finger print, its more efficient and smart way, secure, time efficient than manual conventional system. No proxy is allowed.

Arif [11] this paper aims to record important needed study of on automated attendance systems with timeline. Our peer review interested in the abstract of existing literature on technology, domains and applications.

## 3 Existing System

Most of the existing systems with Face detection and recognition are used to detect the headcount of the classroom and to mark attendance using the detected output. Face revealing and credit system used in security undergoes facial identification to allow authenticated users. The emotion recognition using Audio or speech signal,

it requires a larger computation performance of the GPU when compared to image. The key factor is emotional ability for the next generation robots such as NAO, and Sony AIBO. Twitter uses Sentiment analysis to identify the emotion of the user for every tweet.

## 4 Proposed System

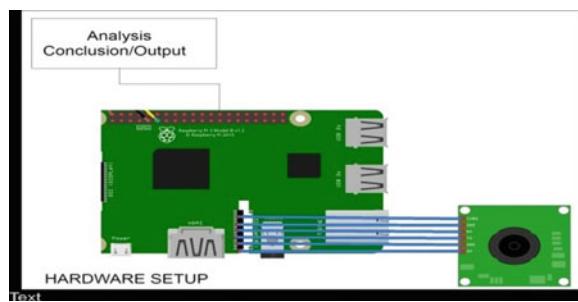
The proposed system is designed to detect the faces and their emotions to determine the satisfaction of the listener. It is the integration of Face Detection, Emotion Classification and Statistical analysis using the Deep Learning. Here the main objective of the algorithm is to implement the CNN alternative for a slow performance hardware constraint and to reduce the number of parameters in training the model. Several methods are brainstormed in finding out proposed system. One method of practical approach is applied to each classroom with some hardware units and data is collected. A compact, low profile projector is attached to the wall of the classroom or a conference hall where the feedback has to be analysed. This projector consists of a camera that will capture the image of the people listening to the lecture and assigns the feedback using their facial expressions. Data mining is done to know the feedback of the students or audience with respect to big data analytics. The result is evaluated by capturing images entire an hour duration and output are distributed under two diverse ways and results are calculated for neutral and happy emotions. This technique will help in valuable way of analyzing any personal feedback without concerning the direct contact method of feedback which sometimes are repetitive. The proposed system is viable to be used in Lecturer hall, conference halls, board meetings, and several cyclic events where the feedback or the survey has to be taken.

## 5 Results and Discussion

### 5.1 *Hardware*

The Overall Hardware setup as shown in Figs. 1 and 2 consists of Raspberry Pi 3B +, Pi Camera, Flat Ribbon cable, Heat sink  $\times$  2, Power adapter. The Pi camera is attached with CSI port in Raspberry using Ribbon Cable and supported by laser cut stand. The recommended Power adapter with the specification of 5VDC is 2.5 A. Heat sinks are essential along with the processors as machine learning processor involves high computational capacity. The processors expel a lot of heat making which might cause some serious issues with its operation in long run. Thus additional heat sink along with Raspberry Pi, which is glued for better support.

**Fig. 1** Block diagram of projector and hardware kit (Wiring hardware module)



**Fig. 2** Block diagram of projector and hardware kit (Output module)



## 5.2 Software

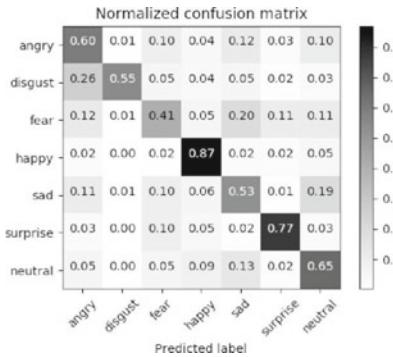
### Real-Time Emotion Detection and Dataset Description

Once the model is trained and the face features detection classifier is defined the next step is to detect the emotions expressed based on these features. For each image frame, the face is extracted and converted to gray scale image with proper processing techniques. By making use of the generated model the Python program calls for the emotion\_Predict() function to find the emotion depicted at the frame. This process is repeated to find the emotion of each faces from the processed image and the information data is log to a csv file. Since that are possibilities of getting false positives at some point, we have introduced the counter measure by aggregating the emotions expressed in the 10 frames and finding the maximum of that list. This ensures only strongly detected emotions from the faces. The Kaggle Facial Expression Recognition Challenge (FER2013) use as reference data set as shown in Table 1 which contains 35,887 pre-cropped,  $48 \times 48$  pixel grayscale images of faces fit in expressions of “angry”, “disgust”, “fear”, “happy”, “sad”, “surprise”, “neutral”.

The Emotion of each face is logged to a csv file. This is done for the mentioned iterations and the graph is plotted to relate the feedback with and the emotions. The emotions are weighted to represent the feedback with a graphical format. The weights are then added and scaled between 0 and 100%. This helps to us to plot the relationship curve for each instance for the required time period.

**Table 1** Imaging of emotions and dataset description

S. No	Emotion	Value
1.	Angry	0–4593 images
2.	Disgust	0–547 images
3.	Fear	0–5121 images
4.	Happy	0–8989 images
5.	Sad	0–6077 images
6.	Surprise	0–4002 images
7.	Neutral	0–6198 images

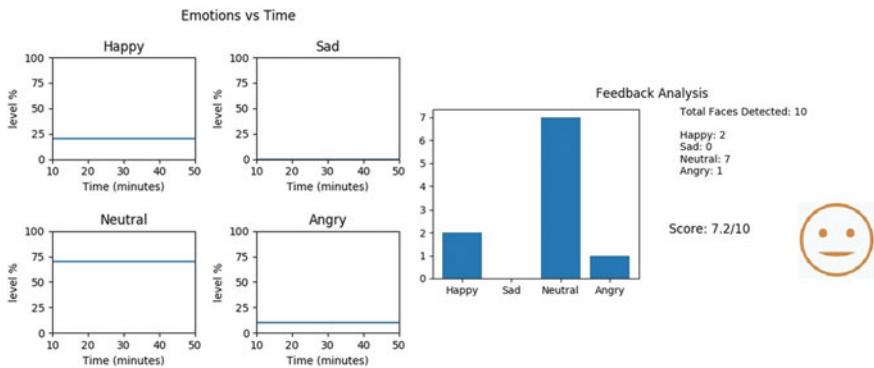
**Fig. 3** Emotion picture dataset 1 and 2

The data mining is again evaluated under different emotion of each faces taken in different moment of time. The processed image and the information data is log to a csv file. By aggregating the measure value the CNN helps in finding new output of score 7.2/10 as neutral as shown in Figs. 3 and 4.

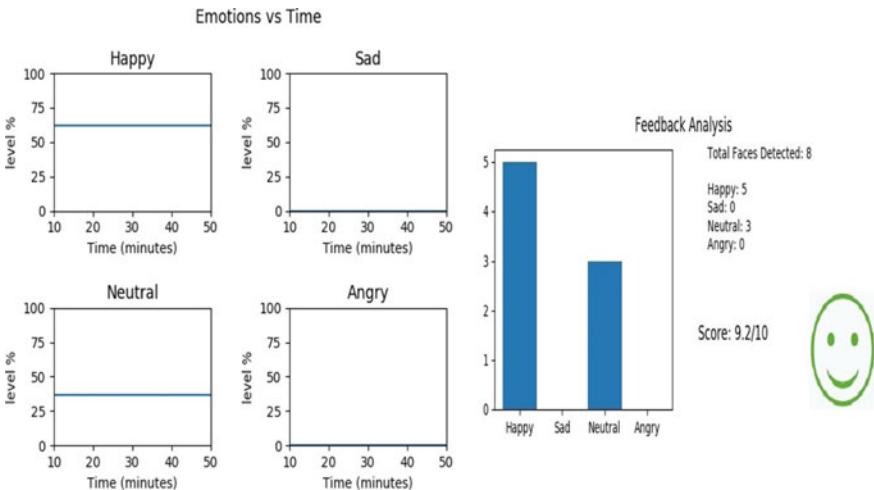
The data mining is evaluated under different emotion of each faces taken in different moment of time. The processed image and the information data is log to a csv file. By aggregating the measure value the CNN helps in finding new output of score 9.2/10 as in Fig. 5.

## 6 Conclusion and Future Work

The project can be adopted in various classrooms and conference hall so that they solve their feedback collection difficulties. The current model of the project is simply



**Fig. 4** FER output for dataset 1



**Fig. 5** FER result for dataset 2

a prototype of a bigger scale version which will be developed in due course of time. This model helps us to understand the basic algorithm and interfacing of the various components used in the project, giving us an idea of how they work on classrooms and this test input images and to determine. The emotions are weighted to symbolize the feedback with a graphical arrangement. The weights are summed up and scaled between 0 and 100%. The relationship curve for each instance for the required time period is plotted. The result are analyzed by capturing images throughout 1 h duration and results are studied under two different ways and results are studied for neutral emotions and happy emotions. This techniques helps in effective method of analyzing any personal feedback without disturbing the direct contact method of feedback which sometimes are boredom. It requires a humongous amount of data

to generate the student's face dataset. We also need to enhance the detection and recognition algorithm to improve the model's accuracy. In the future, we can use powerful hardware modules to scaling this project.

## References

1. S. Shojaeilangari, W.-Y. Yau, K. Nandakumar, J. Li, E.K. Teoh, Robust representation and recognition of facial emotions using extreme sparse learning. *IEEE Trans. Image Process.* **24**(7), 2140–2152
2. L. Kousalya, R. Subhashini, Sentimental analysis for students' feedback using machine learning approach. *Int. Res. J. Eng. Technol. (IRJET)* **06**(04) (2019). p-ISSN: 2395-0072, e-ISSN: 2395-0056. [www.irjet.net](http://www.irjet.net)
3. M.S. Neethu, R. Rajasree, *Sentiment Analysis in Twitter Using Machine Learning Techniques* (IEEE Xplore, 2014). ISBN Information: INSPEC Accession Number: 14063416. <https://doi.org/10.1109/ICCCNT.2013.6726818>
4. D.A.R. Wati, D. Abadiano, *Design of Face Detection and Recognition System for Smart Home Security Application* (IEEE Xplore, 2018). ISBN Information: INSPEC Accession Number: 17577309. <https://doi.org/10.1109/ICITISEE.2017.8285524>
5. H. Zhao, N. Ye, R. Wang, *A Survey on Automatic Emotion Recognition Using Audio Big Data and Deep Learning Architectures* (IEEE Xplore, 2018). ISBN Information: INSPEC Accession Number: 18286740. <https://doi.org/10.1109/BDS/HPSC/IDS18.2018.00039>
6. Y.K. Bhatti, A. Jamil, N. Nida, M.H. Yousaf, S. Viriri, S.A. Velastin, Facial expression recognition of instructor using deep features and extreme learning machine. *Comput. Intell. Neurosci.* hindawi.com/journals/cin/2021/5570870/
7. A. Dzedzickis, A. Kaklauskas, V. Bucinskas, Human emotion recognition: review of sensors and methods. *Sensors* (2020). <https://doi.org/10.3390/s20030592>
8. S. Kuruvayil, S. Palaniswamy, Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *J. Kingaud Univ. Comput. Inf. Sci.* (2021)
9. M.A.H. bin Mohd Nasir, M. Nasir, M.H. bin Asmuni, N. Salleh, S. Misra, *A Review of Student Attendance System Using Near-Field Communication (NFC) Technology International Conference on Computational Science and its Applications, ICCSA 2015: Computational Science and Its Applications* (ICCSA 2015), pp. 738–749. [https://doi.org/10.1007/978-3-319-21410-8\\_56](https://doi.org/10.1007/978-3-319-21410-8_56)
10. R. Naik, M. Mal, S. Koli, A. Karnani, B. Chetwani, Smart attendance management system (SAMs). *Int. J. Emerg. Technol. Innov. Res.* **3**(2), 47–50 (2016). ISSN: 2349-5162. [www.jetir.org](http://www.jetir.org/papers/JETIR1602010.pdf). Available: <http://www.jetir.org/papers/JETIR1602010.pdf>
11. Z.H. Arif, N.S. Ali, N.A. Zakaria, M.N. Al-Mhiqani, Attendance management system for educational sector: critical review. *Int. J. Comput. Sci. Mob. Comput. Mon. J. Comput. Sci. Inf. Technol. IJCSMC* **7**(8), 60–66 (2018). ISSN 2320-088X

# A Study on Automatic Detection of Alzheimer's Disease Using Multimodalities



Ag. Noorul Julaiha and R. Priyatharshini

**Abstract** A leading cause of dementia, Alzheimer's disease (AD) affects the cerebral cortex and worsens with time. It's a debilitating neurological disease that develops progressively over time. The death of brain cells in Alzheimer's disease causes memory loss and cognitive decline. Preventive steps can be taken by the patient to prevent illness. Creating a tracking and reminder system for Alzheimer's patients helps them to complete routine tasks. Alzheimer's disease (AD) and mild cognitive impairment (MCI) have long been diagnosed in patients with neuro-pathological illnesses using neuro imaging. Recent advancement in this area is using multimodal system together with advanced machine learning algorithm to automate the identification and prediction of the progression in Alzheimer disease. This survey focuses on a comprehensive assessment of categorization methodologies and their analytical approaches for predicting Alzheimer disease progression. Also several exhortations for succeeding research in Alzheimer illness have been advised based on the new technology. Along with multimodal diagnosis in the proposed method we will include eye movement tracking, voice analysing and face reading techniques to help in self-evaluation to identify the different stage in the disease.

**Keywords** Alzheimer's disease · Convolutional neural network · Deep learning · K-nearest neighbor (KNN) · Magnetic resonance imaging

## 1 Introduction

With Dementia and memory loss, as well as other chronic and permanent impairments in brain function and daily life independence, are hallmarks of Alzheimer's disease. Alzheimer's disease is a crippling and episodic neurodegenerative disorder.

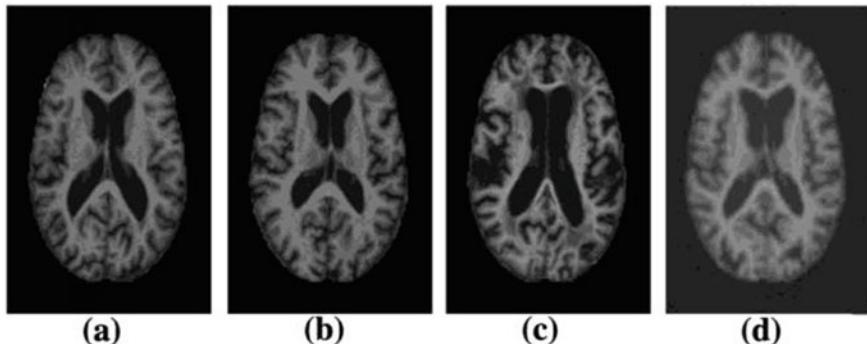
---

Ag. Noorul Julaiha (✉)  
Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India  
e-mail: [ag.nooruljulaiha@gmail.com](mailto:ag.nooruljulaiha@gmail.com)

R. Priyatharshini  
Easwari Engineering College, Chennai, Tamil Nadu, India  
e-mail: [priyatharshini@eec.srmrmp.edu.in](mailto:priyatharshini@eec.srmrmp.edu.in)

People with Alzheimer's disease (AD) are expected to grow from 48 to 153 million by 2050, with significant implications on the economy, healthcare system, and society as a whole. Because of this, no drug exists that can cure AD or halt the illness from progressing completely. There is a greater menace of AD among people with amnestic MCI, because MCI is a step the difference between cognitively healthy ageing and old age and AD. Early diagnosis of AD by MCI screening is critical for appropriate Administrative and clinical methods, as well as for the development of new medications. There are several sorts of techniques available to determine the Alzheimer disease such as Magnetic Resonance Imaging, Electro Encephalogram, Computerized Tomography and Biological marker such as Beta Amyloid Positron Emission Tomography-Computed Tomography PET-CT scan. MRI images are processed for discovering the hallmarks of this degenerative disease with criteria like size, volume, texture, and biomarkers present in plaques and tangles, grey/white matter, thickness, shape, ventricular region, signal noise between neuron Labeling the photos by hand takes time, thus research is focused on automating predictions and early diagnoses of Alzheimer's disease in order to speed up the process. Shape and text features are incorporated to machine learning classifiers to improve patient progression prediction performance.

AD has four stages: early stages of dementia, middle stages, and late-stage dementia, as seen in Fig. 1. Pre-dementia symptoms include forgetfulness and mild cognitive impairment, which mirrors natural ageing. Learning, executive function, and memory problems that are increasingly obvious in people with Alzheimer's disease can produce some linguistic difficulties early on. Reading and writing skills decline, as well as long-term memory, as people progress through the intermediate stage of dementia. Apathy and inability to conduct simple tasks on one's own are common symptoms in Alzheimer's patients in the latter stages of the disease; these individuals eventually become bedridden and die as a result [1, 2].



**Fig. 1** Different brain MRI images from Radiopaedia.org. **a** Pre-dementia. **b** Very mild dementia. **c** Moderate dementia. **d** Late-stage dementia

## 2 Multimodal Data Sources for Alzheimer Disease

The data in the ADNI37 database has been extensively used in studies on Alzheimer's disease ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As part of the study led by ADNI, researchers plan to employ MRIs, PET scans, biological markers, and clinical and neuropsychological testing to track the progression of MCI and early-onset AD. The ADNI is a data set that includes more than 2220 patients imaging, clinical, and genetic information from four different investigations (ADNI1, ADNI2, ADNI GO, and ADNI3). These two studies (AIBL and MIRIAD) use huge datasets to examine ageing in Australia and the brain health of people with Alzheimer's disease, respectively [3].

### 2.1 *Positron Emission Tomography*

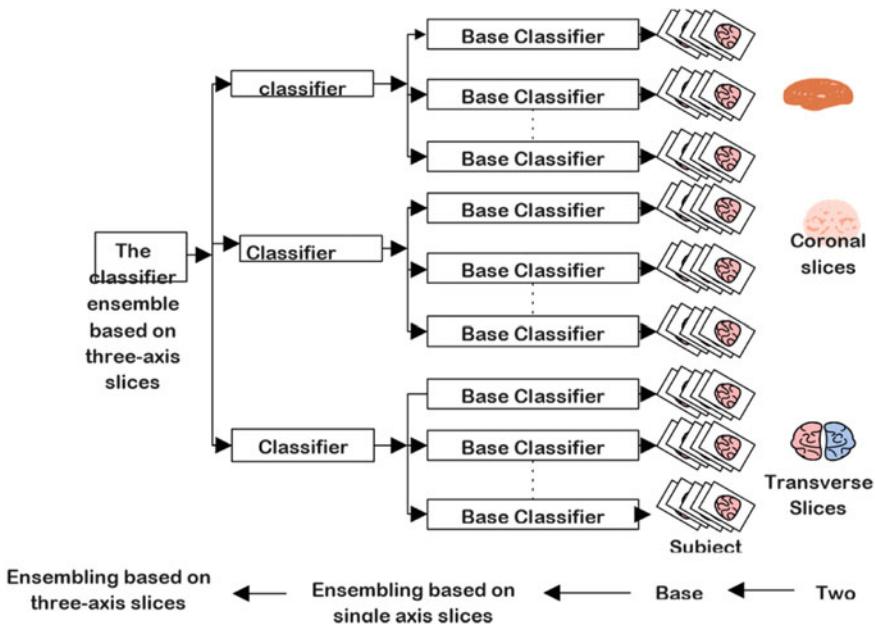
PET scans evaluate various parts of the brain, including the temporal lobe, hippocampus, and frontal lobe, to count the size of each one in order to detect dementia. An MRI scan of the brain reveals any anatomical abnormalities that could contribute to cognitive failure, such as shrinking of the brain or vascular anomalies. MRI scans can help detect if your cognitive impairment is due to anything else, such as a stroke or cancer. PET-CT scans of the brain for a variety of purposes are readily available: fluorescent deoxyglucose is one example of this (FDG) In order to assess brain glucose levels, PET-CT is employed, as well as Beta amyloid proteins and (iii) tau can be measured with PET-CT. Neurofibrillary tangles in nerve cells are detected via PET-CT scans, which look for the protein tau. A positive PET-CT scans for amyloid (increased beta-amyloid levels) could assist diagnose Alzheimer's disease, for example. Tumors, stroke, and traumatic brain injury are all possible causes of dementia that can be ruled out using these diagnostic procedures.

For the most part, MRI is used to look for volume changes in specific locations, which can have an analytical accuracy of up to 87% in the detection of AD. Direct or indirect estimations of mesial temporal lobe atrophy are possible. Measurement of hippocampus or para hippocampal tissue volume loss is used for direct estimation; magnification of para hippocampal fissures is used for indirect estimation. As a precursor of dementia development from MCI, the medial temporal atrophy score is frequently utilized in conjunction with these computations. If you suspect Alzheimer's disease is progressing early on, look for signs of parietal atrophy on the inter hemispheric surface, such as shrinkage of the parietooccipital and cingulate sulcal regions, and a decrease in presumes and cortical area size.

The condition can only be diagnosed with certainty through microscopic study of brain tissue. Mini-Mental State Examination (MMSE) and Mini-Cog test are neuropsychological tests used to evaluate the disease's cognitive deficits. An evaluation of the patient's mood is also done to detect behavioral abnormalities such as worry or depression to eliminate the possibility of co-occurring mental diseases with dementia [4, 5]. Despite the fact that Alzheimer's disease can only be definitively

diagnosed by autopsy using direct examination of the brain matter, new less invasive but equally effective procedures are being developed. Researchers typically use amyloid protein parameters to assess brain size and function based on measurements taken in the cerebrospinal fluid and blood.

Research has concentrated on generating an accurate technical assessment that can be relied on various brain illnesses in order to improve the field because of AD's clinical importance and the wide availability of noninvasive imaging methods in recent years. Researchers have recently developed and implemented Computer Aided Brain Diagnostic (CABD) tools to evaluate the occurrence and severity of Alzheimer's disease. For example, in Fig. 2, the general framework for the application of CABD to Alzheimer's disease (AD) shows that it begins with gathering employing a recognized imaging technique to create mental images of the brain, in order to improve brain imaging, keep employing image preparation techniques, and then moves on to mining vital features by applying a chosen method of feature extraction. Finally, it concludes by applying an evaluation system classifier system to sort the brain image into distinct groups based on prominent feature selection and implementation.



**Fig. 2** Extraction of 2D coronal slices from 3D brain scans of the medial temporal lobe via preprocessing

## 2.2 *Magnetic Resonance Imaging (MRI)*

See Fig. 2.

### 2.2.1 Structural MRI

An extensive body of earlier research has employed brain imaging techniques including structural MRI and CNN models to tell Alzheimer's disease (AD) patients from cognitively normal ones. There are a number of characteristics that limit the applicability of the models produced in earlier studies. This strategy has a few flaws, as you can see. First, the demographics skewed the populations used in research and development and testing. There was another work that created and validated algorithms without using the datasets from the ADNI and Australian Imaging, Biomarker and Lifestyle Flagship Study of Aging using the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset (AIBL). With 3D images as input, the processing needs have risen, making it difficult to construct sufficiently large and deep neural networks. When working with large 3D images, many layers are needed to achieve the desired level of realism. However, creating that many layers can be problematic due to a lack of processing power or GPU memory.

### 2.2.2 F-MRI

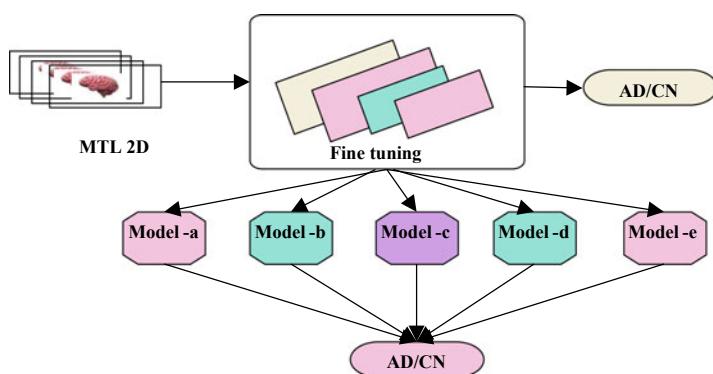
To distinguish between normal and AD cases using T2-weighted MRI images that have been selected, researchers created and applied a CABD method. Pre-processing for the test photographs began with the use of a median filter, which reduces noise in the images. After that, the most critical data from the fMRI scan was extracted using a method known as feature extraction. We carried out a comprehensive comparative analysis of several different method of selecting features including the Curvelet Transform, the Wavelet transforms, such as the Discrete, Empirical and Shearlet types, are examples of wavelet transforms.. The results showed that the CoT, the Curvelet Transform, the DTCWT, and the DWT were the most effective methods of feature selection. This study relied on the  $p$  value and the Student's t-test to choose a subset of attributes for classification. It was finally applied to classify brain MRI data using the k-Nearest Neighbor classifier (KNN).  $256 \times 256$ -by-256 clinical photos (110 normal cases and 55 AD patients) were used to evaluate the CABD system, and the results were impressive (30 normal and 30 AD cases). The CABD system's efficiency was assessed using a three-tiered CV approach. It was decided to evaluate the total performance by calculating the F1 score and the retrieved features' precision, acuity, sensitivity, and specificity.

### 2.2.3 Computerized Tomography (CT)

Bae et al. [6] used CT coronal slices of the medial temporal lobes to develop a CNN-based approach for separating Alzheimer's disease (AD) patients from healthy controls. Our algorithm was tested on two different populations with a wide range of racial and educational backgrounds. However, data from other regions may be useful in the classification of AD. The algorithm may become confused and misclassified with other diseases with similar atrophy patterns if you try to incorporate more places. Since it includes the most important information, we focused on 30 coronal slices to cover the whole hippocampal length. A template-fitting stiff transformation transforms 3D T1weighted MRI images before brain extraction (skull stripping). The skull-stripped template undergoes a second, more severe transformation. After the subject image has been placed in the same spatial location as the template coronal slices have been retrieved from the template-registered output image.

## 2.3 Demographic Data

Demographic factors such as education level can also affect the structural brain alterations linked with Alzheimer's disease. MRI-based AD classification algorithms must be cross validated on populations with varied ethnic and demographic features before their generalizability and practical applicability can be evaluated. According to a recent study, models developed using data from two different populations—one made up mostly of highly educated Caucasians, the other of Asian were directly cross-validated. In terms of generalizability, Database-based algorithms may have an advantage over neuropsychiatric testing because of the inherent ethnic diversity in language, educational attainment, and culture, but neuropsychiatric tests may have limitations (Fig. 3).



**Fig. 3** Developing a classification model for Alzheimer's disease using coronal slices

## 2.4 EEG

To obtain an electroencephalogram (EEG), doctors attach electrodes to the scalp and record electrical brain activity. There are a variety of illnesses for which electroencephalography (EEG) can be used to aid make a diagnosis. Electrophysiological alterations in MCI and AD have been studied in EEG research. Alzheimer's disease can be predicted in persons with mild cognitive impairment using common resting-state EEG metrics including high theta and low beta power as well as a lower mean frequency and spatial alpha frequency characteristics (MCI). Cognition has been linked to both EEG and MRI measurements. Psychometric assessments like the MMSE and the AD Assessment Scale have been linked to EEG beta-1 activity, according to the research. Oscillations in the theta/upper alpha range are indicative of long-term memory activities such as the encoding and retrieval of new information.

## 3 Related Works

When employing selected areas of interest in three dimensions, researchers have been able to avoid this issue. However, the images used were still quite vast. Some of the most complex networks, such as neural networks and neural backbones, use two-dimensional (2D) pictures as inputs. MRI scans as two-dimensional pictures, on the other hand, may be advantageous. Using 2D MRI slices of T1-weighted brain areas, Jong Bin Bae and colleagues created an algorithm for classifying people with Alzheimer's disease based on their racial and ethnic origins as well as demographic characteristics. The results have been cross-validated across the two groups [6].

Memory loss could be detected using deep neural networks in combination with sparse regression models. Various regularisation control settings gave different results when utilising sparse regression models. The ideal weights for integrating various sparse regression models were calculated by CNNs for a hierarchical and non-linear technique to achieve the final prediction values. Slices of white matter and GM were selected for training an ensemble of ConvNet networks based on structural features in MR brain imaging. MR brain imaging reveals localised regions from which 3D patches may be extracted. The patches from each region were then grouped together using the K-Means clustering approach. To get a more accurate final picture classification, the classification results from many local regions were combined [7].

Rajendra Acharya et al. used T2-weighted MRI brain images to develop a CABD method for distinguishing between normal and AD cases. To begin with, the MRI brain scans were scanned in 2D. 66 two-dimensional (2D) test photos with a dimension of  $256 \times 256$  pixels were gathered during this process and examined. After that, a preprocessing approach was used to improve the quality of the test images. A median filtering strategy is utilized in the preprocessing stage to remove test image noise and faults. After that, a preprocessing approach was used to improve the quality of the test images. A median filtering strategy is utilized in the preprocessing stage

to remove test image noise and faults. After that, the brain pictures were mined using methods described in the literature. A student's t-test with a *p*-value of the most essential qualities was found. In the end, a KNN classifier was used to sort the test image into different categories. Finally, the suggested CABD system's accuracy, precision, sensitivity, and specificity percentages were validated [8].

Noninvasive *in vivo* studies of AD-related brain alterations are now possible thanks to advances in brain MRI technology. It has been shown that MRI is useful in many machine learning applications for predicting Alzheimer's disease, such as random forests [9], support vector machines [10], and boost algorithms [11].

For the time being, machine learning approaches include selecting predefined brain ROIs based on well-known Alzheimer's disease MRI indicators (AD). Because reliable MRI biomarkers for Alzheimer's disease are still a work in progress, it's possible that preselected ROIs won't capture everything that could be beneficial in understanding the disease's complexities. Manual selection is time-consuming, labor-intensive, and sensitive to human mistakes [12].

Van der Hiele et al. [13] conducted a cross-sectional investigation to see if brain MRI measurements of anatomy or electroencephalogram (EEG) tests for brain function better predict cognitive performance as people age. It was necessary to pick samples by visual inspection in order to avoid contamination of the EEG due to eye movements, blinking, and muscular activity. Samples were to be 4–8 s long, and they were taken under the standard conditions of eyes closed and open, as well as working memory. Focus-EEG imaging and analysis software was used for frequency analysis (MEGIS software GmdH, Munchen, Germany). The absolute power of the theta band was calculated over a range of frequency bands, from 4 to 13 (Table 1).

## 4 Inferences Made

No single neuroimaging modality can attain ideal accuracy for automatic Alzheimer's disease prediction without combining multiple techniques. The classification issue can only be effectively translated into the clinical realm. The majority of researchers used MRI datasets to identify AD, but there were also studies using EEG, speech, and body motions as data sources to identify the condition as well. The most popular and typical way of detecting Alzheimer's disease is to add features such as shape and texture to photos in order to detect the disease early. Apart from medical diagnosis, behavioral characteristics can also be included for the detection of Alzheimer disease. If advanced technology like genotyping and other cognitive test added with the traditional methods it can even predict the later evolution of Alzheimer disease.

**Table 1** Comparison of the existing classification algorithms

S. No.	Author	Modality	Classifier	Parameters	Output	Merits	Demerits
1	Bae et al. [6]	T1 weighted MRI	2D-CNN	Accuracy	83 82	Takes into account the MTA (medial temporal lobe atrophy) scale	MCI images are not included in the data collection
2	Pan et al. [7]	MRI	CNN-EL	AD versus HC MCIc versus HC MCIc versus MChc	0.84 ± 0.05 0.79 ± 0.04 0.62 ± 0.06	Improved classification precision and resilience	Not applicable to all samples
3	Acharya et al. [8]	MRI	KNN	Accuracy	98.48	There are less features used	Classification accuracy is lower
4	Feng et al. [14]	MRI	2D-CNN 3D-CNN 3D-CNNSVM	NC versus MCI	82.57 ± 7.35 89.76 ± 8.67 95.74 ± 2.31	There will be no manual feature extraction	The level of automatic diagnosis of Alzheimer's disease is not up to that of radiologists
5	Khatun et al. [15]	EEG	SVM	ROC	91	MCI prognosis based on EEG	There is no implementation in real time
6	Varatharajan et al. [16]	Wearable sensor for foot movement monitoring	Dynamic time warping algorithm	Sensitivity Specificity	95.9 94	The walking pattern is used to diagnose Alzheimer's disease	Various physiological cues must be considered
7	Mahmood et al. [17]	MRI	Principal component analysis	Accuracy	89.92	Automated classification of Alzheimer's disease at various levels	Different parts of MR pictures will be examined separately

(continued)

**Table 1** (continued)

S. No.	Author	Modality	Classifier	Parameters	Output	Merits	Demerits
8	Odusami et al. [18]	fMRI	Finetuned ResNet18 network	EMCI versus AD LMCI versus AD MCI versus EMC	99.99 99.95 99.95	EMCI and LMCI are intermediate stages of MCI that can be identified	Overfitting was produced
9	Ali et al. [19]	MRI	TANNN technique	Accuracy	99.2	Image mining is used to identify Alzheimer's disease	The type of Alzheimer's disease is unknown
10	Kanghan et al. [20]	MRI	CNN and TL	Accuracy	86.60	The most discriminating bio indicators associated to AD/MCI are displayed	The number of subjects employed in the training and testing phases was limited
11	Lebedev et al. [21]	MRI	Random forest ensembles	AD/HC classification	83.3/81.3	Examining the most critical aspects that affect classification performance	The impact of suspected diagnostic data mislabeling on the results

## 5 Conclusion

Alzheimer's disease can be stopped or slowed down if caught early enough. This study provides an overview of numerous strategies that have been previously proposed by researchers the incident as it stands of automatic detection of AD and also spoke about the severity of the condition. Machine learning approaches like CNN SVM, Random Forest and KNN are thoroughly reviewed in this paper. By combining MRI and EEG modalities, we can better detect and track disease progress. Research into Alzheimer's disease focuses on finding a way to detect the disease early so that it doesn't progress to the next stage. To increase the accuracy of illness prediction, the model's parameters are expanding. The objective is to keep the death rate of Alzheimer patients low before the disease progresses to a severe stage. This prediction can be improved in the future so that Alzheimer's disease can be classified as AD or MCI, for example. Also, rather than viewing the MR image as a whole, we may examine distinct regions to determine Alzheimer's disease affects different areas of the brain. Besides that, we're interested in the ventricle-to-hippocampus expansion ratio loss as a way to define the course of Alzheimer's disease. Our capacity to extract reliable biomarkers of neurodegenerative disorders will be greatly enhanced by combining many techniques (such as CNN, RF, KNN, Deep-learning, and SVM), as well as multiple modalities (MRI, EEG, and PET).

## References

1. G.A. Malik, N.P. Robertson, Treatments in Alzheimer's disease. *J. Neurol.* **264**(2), 416–418 (2017)
2. Editorial, The three stages of Alzheimer's disease. *Lancet* **377**(9776), 1465 (2011). Refer <https://pubmed.ncbi.nlm.nih.gov/21531256/>
3. J. Venugopalan et al., Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci. Rep. Nat. Portfolio* (2021)
4. J.F. Norffray, J.M. Provenzale, Alzheimer's disease: neuro pathologic findings and recent advances in imaging. *Am. J. Roentgenol.* **182**(1), 3–13 (2004)
5. L.J. Whalley, Spatial distribution and secular trends in the epidemiology of Alzheimer's disease. *Neuroimag. Clin. North Am.* **22**(1):1–10 (2012)
6. J.B. Bae et al., Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci. Rep. Nat. Res.* (2020)
7. D. Pan, et al., Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning. *Front. Neurosci.* **14** (2020)
8. U.R. Acharya et al., Automated detection of Alzheimer's disease using brain MRI Images—a study with various feature extraction techniques. *J. Med. Syst.* (2019)
9. J.M. Mateos-Pérez, M. Dadar, M. Lacalle-Auriolles, Y. Iturria-Medina, Y. Zeighami, A.C. Evans, Structural neuroimaging as clinical predictor: a review of machine learning applications. *NeuroImage Clin.* **20**, 506–522 (2018)
10. E.E. Tripoliti, D.I. Fotiadis, M. Argyropoulou, A supervised method to assist the diagnosis and monitor progression of Alzheimer's disease using data from an fMRI experiment. *Artif. Intell. Med.* **53**, 35–45 (2011)

11. K. Leemput, F. Van Maes, D. Vandermeulen, P. Suetens, Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imaging* **18**, 897–908 (2002)
12. C. Hinrichs, V. Singh, L. Mukherjee, G. Xu, M.K. Chung, S.C. Johnson, Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* **48**, 138–149 (2009)
13. K. Van der Hiele et al., EEG and MRI correlates of mild cognitive impairment and Alzheimer's disease. *Neurobiol. Aging*, 1322–1329 (2007)
14. W. Feng et al., Automated MRI-based deep learning model for detection of Alzheimer's disease process. *Int. J. Neural Syst.* **30**(6) (2020)
15. S. Khatun et al., A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses. *Neural Syst. Rehabil. Eng.* (2018)
16. R. Varatharajan et al., Wearable sensor devices for early detection of Alzheimer disease using dynamic time warping algorithm, in *Cluster Computing* (2017)
17. R. Mahmood et al., Automatic detection and classification of Alzheimer's disease from MRI scans using principal component analysis and artificial neural networks, in *20th International Conference on Systems, Signals and Image Processing (IWSSIP)* (2013)
18. M. Odusami et al., *Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images Using a Finetuned ResNet18 Network, Diagnostics* (MDPI, 2021)
19. E.M. Ali et al., Automatic detection and classification of Alzheimer's disease from MRI using TANNN. *Int. J. Comput. Appl.* **148**(9) (2016)
20. K. Oh et al., Classification and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning. *Sci. Rep. Nat. Res.* (2019)
21. A.V. Lebedev et al., Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage Clin.*, 115–125 (2014)

# IoT Enabled Smart Tourism (IoTeST): Tourism Service Dimensions



Shivani Sharma and O. P. Rishi

**Abstract** Smart tourism is a social, cultural, and economic phenomenon where people are displaced to tourist destinations for personal or professional reasons. The main purpose of travel and tourism is cultural immersion. In the early 1990s, cultural tourism became one of the alternative forms of tourism to mass tourism. Tourists take a proactive approach and actively create experiences while traveling. Hence, the heterogeneous preferences of the tourists have drawn the attention of the travel destinations, and the destination cities are striving to provide high-quality and personalized travel experiences. The tourism industry is pursuing successful etourism strategies through website development, e-commerce, and related governance. The convergence of ICT and tourism describes the transformation of tourism through technology. Smart tourism means tourism using advanced information technologies to make travel experiences more enriching, efficient, and sustainable. In this sense, smart tourism is a social phenomenon that arises from integrating ICT into the tourist experience. In addition, smart technology like IoT has significantly changed the way tourists make various travel decisions, such as transportation, accommodation, and available activities in the desired tourist destination. It changes the expectations and behavior of tourists. This change in the IT-enabled tourism segment has been termed as “IoT enabled smart tourism” (IoTeST). It is streamlining all procedures and processes of hotels booking, airline reservations, train booking, and other tour & travel firms by connecting smart gadgets. In the current scenario, along with wi-fi, NFC, smartphones with internet connectivity, RFID, and clouds, IoT has become the core heart of smart tourism. As soon as tourists start exploring the tourist destination’s places, IoT technology starts streamlining the entire tourism process and optimizing the entire travel experience.

**Keywords** IOT · NFC · RFID · Disruptive technologies · Seamless travel · Smart tourism

---

S. Sharma (✉) · O. P. Rishi

Department of Computer Science & Informatics, University of Kota, Kota, India  
e-mail: [shivanijisharma1@gmail.com](mailto:shivanijisharma1@gmail.com)

O. P. Rishi  
e-mail: [dr oprishi@uok.ac.in](mailto:dr oprishi@uok.ac.in)

## 1 Introduction

The tourism perspective will completely change with the help of IoT technology soon. It would create excellent opportunities and challenges for the tourism industry. With the continuous flow of information, people will understand a lot about the world, and after all, travel and tourism businesses. With the huge amount of information and variety of sources relating to tourism, it is very important to use of this information. Collecting and categorizing the process for the raw input data to improve the quality of tourist services and their manageability [1]. The key mandate for the travel and tourism industry is to attract more and more tourists from entirely different parts of the world. However, lack of information, insecurity, fraud, lack of adequate means of transport, and communication are the most critical problems. In addition, security considerations in tourist locations have limited the independent mobility of tourists [2]. They were paying for goods while traveling. Payment gateways removed all the hurdles as well as to increased security during on-site currency conversion for the purchase. IoT enabled Smart Tourism (IoTeST) will simplify the entire travel and tourism process. Hanging from lampposts, sitting in heritage parks or gardens, strapped to the roofs of a tourist bus, and rooted in parking embankments covering the asphalt. IoT sensors help tourists find and apprehend data about anything and everything that they have a time looking through the lens of their smartphone [3]. The IoT sensors collect information about tourists' behaviors, feedback, and responses to see tourist trends and optimize the excellent travel experience. Based on the established data, the tourists will serve the appropriate services [4, 5]. Disruptive technologies like IoT, artificial intelligence, and intelligent agents are essential for increasing efficiency and productivity in the tourism sector. These technologies have the capability to gather data from different sources and transfer it to cloud platforms, where this data is analyzed and used for better decision-making processes for the future through big data analytics. This is the first step in emphasizing the importance of IoT in smart cities as well as smart tourism [4, 6]. When smart cities and smart tourism concepts have been automatically transferred to tourist destinations, we call it smart destinations. Here, it is very significant to emphasize that Smart destinations are just smart cities with an additional level of complexity of technology management. In this way, all smart city applications which are utilizing IoT and different IoT based technologies are directly applicable for the identification of smart destinations and tourism activities [5].

Furthermore, suppose smart cities aim to improve the quality of life of their inhabitants. Then, Smart tourist destinations automatically improve the quality of life of each person who spends their time in these destinations [7]. The whole process would not be complete without turning tourists themselves into smart tourists. For the same, these smart tourists need to change their travel habits like they travel more and more and use smart gadgets and mobile devices, particularly smartphones with intelligent travel support applications. Smart destinations require various sensors to capture real-time data and intelligent applications to integrate and analyze this data for various stakeholders. In this procedure, the IoT plays a crucial role in capturing real-time



**Fig. 1** Component of IoT enabled smart tourism

data to improve tourism experiences, managing the destination more proficiently, and providing information exchange channels (Fig. 1).

## 2 Travel and Tourism

The travel and tourism industry comprises various stakeholders like local residents of destination places, business associates, government, tour operators, transporters, and tourists. Personalized tourist experiences are at the top in the travel and tourism industry. For the personalized tourist experience, IoT plays a critical role in the improvements and industry. However, this is only possible if the on-site IoT infrastructure allows devices that provide the real-time information needed to make real-time decisions. Real-time information helps to create simulations to predict scenarios about the destination. Therefore, tourist destinations need to set up complete systems for collecting, analyzing, and disseminating information between all the actors included in the tourist destination. The intent is to create a smart destination that the systems collect, analyse, and disseminate information between all tourism sector actors and implement an excellent tourist strategy that will increase its value using IoT innovation [8]. This process of developing smart destinations, smart transports, and smart gadgets will help better utilize existing resources. With the help of these excellent resources will be identified and created for increasing competitiveness. It will also improve the efficiency of tourism marketing, production processes, and the use of renewable energy sources [9, 10].

## 3 Internet of Things and Smart Tourism

Internet of Things is (IoT) a network of properties through which collect, share or exchange the data between physical devices or physical objects. These physical

objects or devices may be deployed in vehicles, buildings, and other things. These physical devices or objects are integrated or embedded with software, sensors, and network properties that enable these objects to gather and communicate information between them [2]. The IoT can make it easier to collect immense amounts of information while still capturing and analyzing it. It requires the right technology and the right technocrats and decision-makers to draw actionable decisions from the information. Future transformation in the IoT technology may cause further disruption to the tourism industry [11]. The architecture of IoT varies from problem to problem and corresponding solutions. Its architecture depends on the solution we intend to build hence because of this IoT enabled smart tourism may cause further disruption. However, the IoT as a technology for the tourism industry mainly consists of four main architectural components: sensors, devices, gateway, cloud, and tourists [5]. One of the furthermost overcritical characteristics of IoT-enabled tourism is the ability to know the likes and dislikes of the tourist and provide tailor-made personalized services during the travel, during the stay in a hotel, and travel in the aircraft or train. In this way, the tour operator or travel associate makes it possible to personalize services for the tourist to meet their expectations. For example, suppose a tour operator or associate has this information that one of our customers likes vegan food. In that case, they can send this information to culinary workshops, and the customer can get the specialized food as he or she visits the restaurant [12].

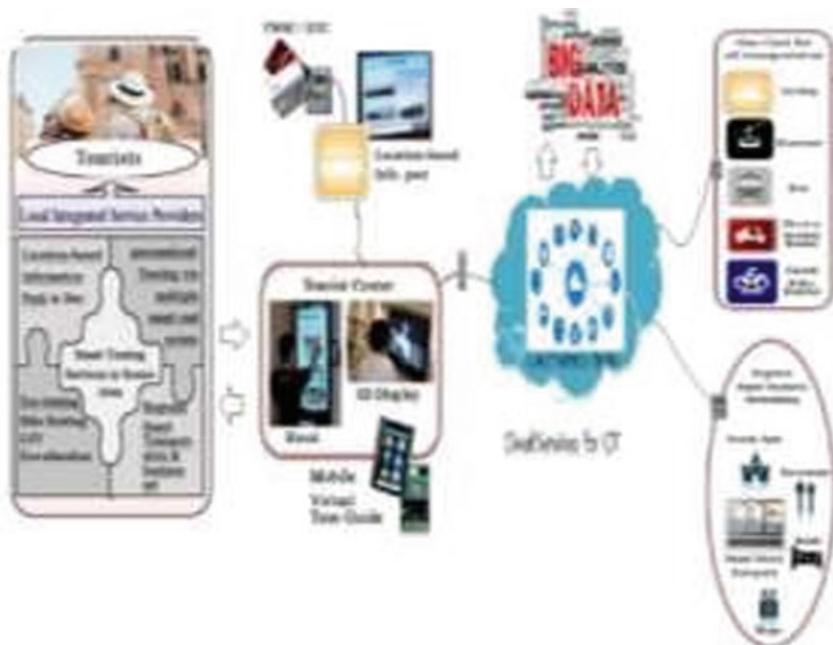
## 4 IoT Enabled Smart Tourism (IoTeST)

The IoT enabled smart tourism (IoTeST) will impact travel bookings, baggage handling, and services in hotel rooms, availability services at airports and train stations, in-flight services, and tourist services [3]. The different technologies available to provide good tourism services are web technology, mobile technology, cloud computing, IoT (IoT), and big data analytics. Simple devices like our cell phones are equipped with a camera, microphone, sensors to detect movement, etc. the information collected by these smartphones is often processed on servers and then used for Smart tourism and Smart livelihoods [5]. Smart livelihoods require Smart homes, smart offices, smart transportation, smart cities, and Smart working conditions. Its backbone is because it provides real-time information. Therefore, smart devices such as smartphones, Smart information kiosks equipped with sensor cameras, microphones, and touchpads are very important tools to provide real-time data to tourists to plan trips. Thus, the impact of IoTeST in the tourism industries are as follows (Figs. 2, 3 and 4):

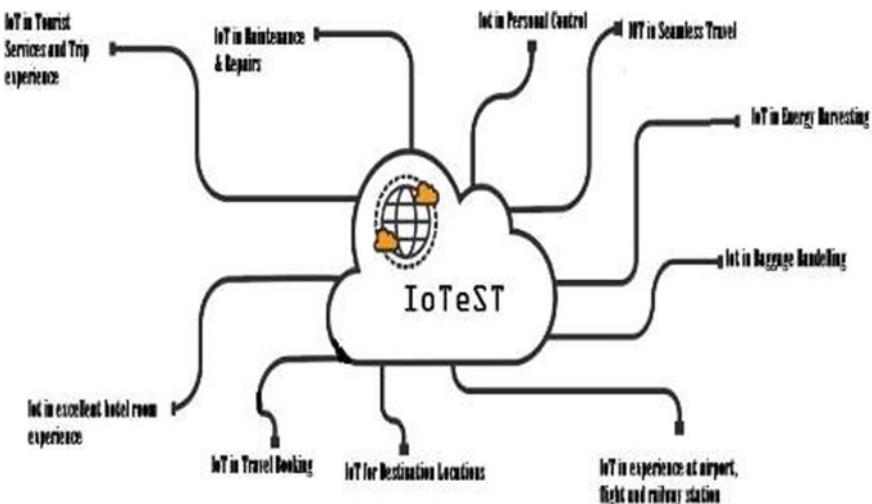
- Personal control
- Unified travel
- Location Information
- Travel bookings



**Fig. 2** Architectural component of IoTeST [5]



**Fig. 3** Overall architecture of IoTeST [5]



**Fig. 4** Impact dimensions of IoTeST in the tourism industries

- Experiences at the airport, flight and the station:
- Baggage handling
- Hotel Room Experience
- Smart Energy Saving
- Maintenance and Repairs
- Tourist Services and trip experience
- Services on future trips.

#### 4.1 *IoT in Personal Control*

The foremost extensive uses of IoT technology for enabling a better degree of personalization in hotels and flights for services. This is often achieved by permitting clientele tabs or smartphones to regulate additional appliances or services by a centralized device. Suppose in the hotels, equipment or appliances like cooling, lighting, heating, television, washing machine, etc. are internet enabled. In that case, customers or tourists can operate these appliances far from where they are. They can manage these appliances remotely and select a selected temperature and light according to their need and choice. Similar IoT technology can be used in flights for regulating seat temperatures, light on the seat [10].

## ***4.2 IoT in Unified Travel***

After personal control, another exceptional use for IoT in all areas of the travel industry is streamlining customers' experiences for unified travel. For example, suppose any passenger is keen to know about their bags at the airport. Then with the help of IoT tags, passengers can identify their bags are nearby with the help of sensors on their smartphones in the fastest manner. As soon as customers reach the hotel, the arrival method is created seamlessly, which means the hotel sends the electronic key or pin to the guest on their smartphone. In this way, the customer can check-in automatically in the hotel without the help of the front desk. After checking in the hotel, the customer may switch on the light or set the room temperature with the help of their smartphone. As soon as customers enter the restaurant, sensors warn the restaurant house staff about the guest. Once guests arrive, they automatically by assigning and send them to the correct table number by assigning the correct table based on desired services available [4]. At each level, all the services are available to all customers. This is a unified travel experience for the tourist or customer.

## ***4.3 IoT in Destination Location Information***

Companies running their business in an expedition, heritage, or destination tourism can also use the IoT to send location-specific data to the clients and acquire valuable data from the destination places. With the help of sensors and smartphones, the tour operator or tour associate can collect and send the destination-specific data or information to the tourist wherein they are. For instance, messages about neighborhood points of interest, the shortest distance, and the required time to visit can be despatched to the clients. In other instances, while clients are free, messages from nearby public shopping offers may also be despatched. In addition, messages and alarms can be sent to the clients about free occupancy in hotels and lodges with day-to-day tariffs. Suppose hotels and lodges are entirely occupied, and no free occupancy exists. In that case, an alarm message can be sent to the clients to choose the alternative destination instead of the existing one [7, 13].

## ***4.4 IoT in Travel Bookings***

With the help of IoT enabled tourist system, any tourist can access any information related to their travel bookings from a single connection point. Within this system, all the tourist-related tools and resources are well interconnected. Anyone can get the desired information related to booking and schedule on this interconnected travel platform. The travel agent can easily manage the tourist reservations in hotels, flights, trains, and local transport like buses and taxies and similarly track

and report the tourist information. Hence, the booking of hotels, flights, local transports, and traveling, managing the traveler's entire trip becomes fully connected and optimized with travel plan with the help of IoT enabled tourist system [7]. Excellent connection gives greater control to all the tourists as well as the people who are associated with the tourists and gives them better visibility as they need to elevate their travel plans. Travel assistant chatbots with the help of IoT can also increase the travel business's productivity. Even some tour operators are using such chatbots in their businesses, and the influence of these chatbots increased in recent years for the travel and tourism business [4].

#### ***4.5 IoT in Service Experiences at the Airport, Flight, and the Railway Station***

In the travel and tourist system, the airports are the most prominent and important points of entry at local and international travel destinations. At the airports need to place the IoT application for the help of tourists and travelers. With the IoT-enabled devices and applications, every traveler and tourist can acquire the required personalized information related to the local and global destinations, connecting transportation and infrastructures straightforward. Hence, airports are hubs connecting the entire world, directly affecting safety and security. Air terminal chamber global affirms IoT potential in air terminals and the flying exchange through the functional upgrades and information trade among the partners [1, 14]. Moreover, the info sharing among the cooperative stakeholders can alter them to create higher selections resulting in better client service at traveler screenings. Technology is making new opportunities at an occasional value for the shipping trade. It can remodel several novel techniques that cherish improved property to airplanes and baggage chase etcetera accessible globally and straightforwardly deploy [14]. IoT permits travelers that when they enter the air terminal and the air terminal, they can identify the vacationer at registration and send him back without lining for visa control. Indeed, even travelers can explore a packed air terminal and train stations utilizing IoT. The IoT-based excellent doors at the air terminal and the rail line station recognize traveler-based biometric acknowledgment (like facial and retinal), decreasing check-in time. Travelers should initially enroll their biometric information (a one-time process), yet the entryway opens consequently later acknowledgment. IoT-empowered security controls speak with one another and send a warning to a tourist cell phone with assessed stand-by times and delays [4, 6]. IoT apps can make it easier for tourists to connect to the many devices they carry with them. If any tourist wants to change flight or have connecting flights, real-time notification information can be obtained on the flight status. The tourist can also get directions to the airport boarding gate, instructions on onboarding an aircraft, and information related to security checks carried out [6].

While tourists fly in the aircraft, the IoT sensors automatically coordinate with other sensors and continuously monitor, measure anxiety level, pulse, important bodily functions, affiliation level, travelers, and allow cabin crew to provide comfort. In addition, IoT applications enable to work in the aircraft on the general cost and optimized fuel utilization, considering energy costs. IoT enable system also encourages when and where to refuel and save a lot of fuel.

#### ***4.6 IoT for Baggage Handling at Airports***

At the time of traveling, at the airport, the huge number of bags are mishandled every year, and improper baggage handling issues cause inconvenience to customers [15]. It also leads to high prices for the travel industry. To remedy this problem, the Internet of Things can help by creating luggage tags, and because of this, our bags will be less lost during the outing, which will be conceivable due to things arrangements dependent on sensors. We will get a pop-up message when our baggage is stacked, when it arrives, and where it will be transferred. Also, perform a speedy baggage hand-off process by the IOT enabled system at self-service points without any manual interaction. IoT enabled systems can make it easier to simplify day-to-day operations at airports and are ready to communicate their delivery time and the real location of the luggage with enhanced security. For this, the sensors attached as a tag with to luggage and it be will be useful for travelers to identify the location and distance of the baggage from the traveller.

Suppose ant of the traveller or tourists loses their baggage during the travel. In that case, these sensor tags allow airline companies to find out and deliver the lost baggage quickly to the concerned customer [5, 6].

#### ***4.7 IoT in Excellent Hotel Room Experience***

The appearance of new advancements in IoT for the field of tourism has not gone unrecognized, and it was presumably the area that has benefited the most from the computerized age. However, there are still not many explicit IoT applications in this industry. IoT applications show incredible potential in the travel industry area. By introducing various sorts of sensors in hotel rooms, these applications platform allows monitoring of sensor usage at any time. It may seem like this is just one more home automation project. However, the great potential of these software applications lies in two viewpoints: from one perspective, the hotel visitors who can utilize these App or a program to change or manage his/her own room temperature, turn on the TV, select the appropriate shade of the LED lights, and so on. It also permits the guest to modify his/her room and partake in an Aunic experience [6]. But the real potential of smart tourism applications comes from when hoteliers will be able to find out how long guests will stay in their rooms, the schedule of the guest and

exactly what guests expect from them, and more. The room entry of the guest in the room by smartphone as a key and the room occupancy can be detected with Smart sensors and Bluetooth [4, 6]. Suppose we add to all these innovations to make the system intelligent. In that case, we make a goliath step towards streamlining energy utilization without influencing the solace and prosperity of clients [5]. “While staying in hotels, IoT devices and applications may come up with offers that allow tourists or guests to open curtains, control indoor thermostats, or control TVs. For example, the connected thermostat can regulate the incoming and outgoing room temperature and eliminate the costs of cooling or heating vacant rooms.

#### ***4.8 IoT in Energy Harvesting***

The IoT can provide significant help for energy harvesting by modifying the personalization use of equipment. It may also supply monetary benefits through IoT based remote machine-controlled energy saving tools. For instance, internet-enabled IoT devices and sensors are deployed for temperature management, which means heating is simply used once when needed. Similarly, an analogous principle may also apply to light-weighting as per the need of the light. Numerous hotels already use this IoT-enabled technology to regulate lights turned on and off. Various sensors automatically observe the amount of natural light within the room and turn on and off the room lamp accordingly. This means less energy is wasted, and powerful light-weighting is essentially utilized once normal light isn't enough [6].

#### ***4.9 IoT in Maintenance and Repairs***

The IoT is again useful for maintenance and repair purposes as sensors and associated application tools are installed to give important information on time regarding their present standing and working request. Before the failure of any equipment or device, these sensors give the appropriate information to the concern for the maintenance purpose, so the concern can take the appropriate action before the failure. This kind of IoT-enabled technology will be very important in the travel and tourism sector, allowing virtual gadgets to be fixed or supplanted before they quit working [4].

#### ***4.10 IoT in Tourist Services and Trip Experience***

Tourism services and IoT travel experience will help and serve travellers by giving important data about their visit. Data relating to the tourist likes and dislikes will facilitate exceptional services. Before issues from tourists arise, they will be retained.

It becomes efficient service management to please every tourist, leading to total loyalty [16].

#### **4.11 IoT in Services on Future Trips**

During their excursion to the city or travel, every tourist will want to find all the information and notices related to their travel (travel data) accessible on their cell phones. There will be sensors in all traveling spots of objective, inside the tops of trains and transports, stops, gardens and parking areas, and so on. These many sensors will gather the data based on travelers' habits and responses. This information data is used for totally different fascination, determining patterns, and improving the tourist experience [9].

### **5 Conclusion**

The IoT Platform is very useful in the era of digital Life. It is the shortcut to doing Things Effectively and Efficiently for Both Consumers and businesses [13]. The boom in IoT technology will drive the future of the hospitality industry; it will offer a competitive advantage in the market and thanks to the interconnection of gadgets (sensors, actuators, ID names, phones, and so forth) through the Internet. The IoT is an immanent part of the industry; it is no longer just a concept, and more insights are expanding as IoT empowers cycles, information, and results.

### **References**

1. A. Dalli, S. Bri, Design of electronic ticket system for smart tourism. *Adv. Comput.* **7**(1), 11–14 (2017)
2. M. Pustišek, L. Mali, A. Kos, Scopes in internet of things for active tourism and leisure. ERK, Portorož, A (2014), pp. 85–88
3. M. Nitti, V. Pilloni, D. Giusto, V. Popescu, IoT architecture for a sustainable tourism application in a smart city environment. *Hindawi Mob. Inform. Syst.* 1–9, Article ID 9201640 (2017). <https://doi.org/10.1155/2017/9201640>
4. A. Gomez-Oliva, J. Alvarado-Uribe, M.C. Parra-Meroño, A.J. Jara, Transforming communication channels to the co-creation and diffusion of intangible heritage in smart tourism destination: creation and testing in Ceutí (Spain). *Sustainability* **11**(14), 3848 (2019)
5. S. Sharma, O.P. Rishi, Sharma Anukrati; “IoT enabled smart tourism”(IoTeST): Shaping the Future of Tourism; Rising Threats in Expert Applications and Solutions Springer Nature Singapore Pte Ltd. 2021; pp 569–576.
6. M. Dolores Ordóñez, A. Gómez, M. Ruiz, J.M. Ortells, H. Niemi-Hugaerts, C. Juiz, A. Jara, T.A. Butler, IoT technologies and applications in tourism and travel industries (2020), pp. 341–360 (downloaded from Semantic Scholar; Corpus ID: 231404089)

7. C. Lamsfus, D. Martín, A. Alzua-Sorabal, E. Torres-Manzanera, Smart tourism destinations: an extended conception of smart cities focusing on human mobility. *Inform. Commun. Technol. Tourism* **2**(3), 363–375 (2014)
8. S. Anukrati, O.P. Rishi, A study on e-marketing and e-commerce for tourism development in Hadoti Region of Rajasthan, in *Proceedings of International Conference on Information and Communication Technology for Intelligent System* (2018)
9. K. Kaur, R. Kaur, Internet of things to promote tourism: an insight into smart tourism. *Int. J. Recent Trends Eng. Res.* **02**(04), 357–362 (2016)
10. R. Roshan, O.P. Rishi, Smart solar street light using WiFi, IR motion sensor and LDR for the smart city. *Int. J. Innov. Technol. Exploring Eng. (IJITEE)*, 75–79 (2019)
11. R. Roshan, A. Sharma, O.P. Rishi, IoT platform for smart city: a global survey, in *Emerging Trends in Expert Applications and Security* (Springer, Singapore, 2019), pp. 197–202
12. O.P. Rishi, S. Anukrati, A connected world through IoT: the future of travel tourism. *ICT for Hospitality tourism, and Travel industry*. Cambridge Scholars' Publishing (2018), pp. 1–22
13. D. Buhalis, A. Amarangana, Smart tourism destinations enhancing tourism experience through, in *Information and Communication Technologies in Tourism*. Cham (2015)
14. A.K. Tripathy, P.K. Tripathy, N.K. Ray, S.P. Mohanty, iTour: the future of smart tourism. *IEEE Consumer Electron. Mag.* 1–5 (2018)
15. M. Lacinak, J. Ristve, Smart city, safety and security. *Procedia Eng.* **192**(Supplement C), 522–527 (2017)
16. R. Amjed, M. Sadik Croock, Internet of thing based tourism system: survey and proposed solutions. *Int. J. Adv. Res. Computer Eng. Technol. (IJARCET)* **7**(2), 118–122 (2018)

# An Innovative Approach Towards Designing Efficient K-Nearest Neighbour Algorithm



D. Abhyankar and Maya Ingle

**Abstract** k-Nearest Neighbour (kNN) Algorithm is extensively used in applications of machine learning since it is easy to implement and leads to high performance. Although literature covers machine learning aspects of kNN algorithm, there is a dearth of work in terms of improving its execution efficiency. In order to classify an object X using kNN, algorithm needs to compute the distances between object X and all the N training objects. Computing N distances is costly in terms of running time. In order to reduce this cost, the data structure such as K-D tree leads to minimize the running time, but increases the software complexity and maintenance cost at the same time. Thus, there remains a scope to overcome these challenges. An innovative and efficient approach in kNN is designed viz. easy to implement and maintain. This approach will be highly useful in developing, classifying and implementing real-time systems based on machine learning.

**Keywords** K-means · Classification · Machine learning

---

D. Abhyankar (✉) · M. Ingle  
SCS&IT, DAVV, Indore, India  
e-mail: [deepak.scsit@gmail.com](mailto:deepak.scsit@gmail.com)

## 1 Introduction

Machine learning reveals the way to build computer systems that are intelligent and improve through experience automatically [1]. K-Nearest Neighbour (kNN) algorithm is extensively used in applications of machine learning since it is easy to implement with high performance [2]. On the other hand, an intriguing study suggests ensemble of BPSO, Adaboost along with kNN [3]. Also, there is a study of kNN that covers machine learning aspects of the algorithm [4]. Literature also reveals another aspect of kNN referring algorithms used for classification of an object. It needs to compute the distances between object X and all the training objects. If there are N training objects, then N distances will have to be computed. Computing N distances is costly in terms of running time. In order to reduce this cost, literature suggests using data structure such as K-D tree. The problem with K-D tree approach is its software complexity [5–8]. Instead of implementing such a complex approach, this study approaches the problem in a way that is easy to implement, easy to maintain, and easy to prove efficient. To reduce the cost of computing N distances, this study engineered an elegant algorithm.

However, there is a scope to explore algorithmic aspect of kNN.

Algorithmic approach of kNN has been used widely in recent real-time applications. In Sect. 2, the details of existing algorithm OneNN and its ramifications have been focused. Section 3 covers an innovative approach to design an efficient algorithm. In Sect. 4, mathematical analysis in the form of probability estimates have been covered. In Sect. 5 key results have been emphasized succinctly. Section 6 provides a conclusion of this study.

## 2 Existing OneNN Algorithm

This algorithm is an instance of kNN for  $k = 1$  and used to classify the object. A succinct coverage of the existing algorithm is presented in this section.

## 2.1 Informal Description

Let us assume that  $N$  as number of training objects, each having  $d$  dimensions. Here, L2 distances are calculated to find the nearest object. OneNN is a key function of this algorithm that calls L2. The distance between objects  $X$  and  $Z$  is given by  $L2Distance(X, Z) = \sqrt{((X_1 - Z_1)^2 + (X_2 - Z_2)^2 + \dots + (X_d - Z_d)^2)}$ .

## 2.2 Formal Description

*/\* Algorithm: OneNN \*/*

```

const int N = 1000;  const int d = 100;
Struct TrainingData{
    int x[N][d];
};

int OneNN(TrainingData T, int N, int* Z){
    int CurrentMinimum = INT_MAX;
    int j = 0;
    int NearestObject = 0;
    while(j < N){
        int t = L2((&(T.x[j][0]),Z,d));
        if(t < CurrentMinimum){
            CurrentMinimum = t;
            NearestObject = j;
        }
        j++;
    }
    return NearestObject;
}

double L2(int* X, int * Z, int d){
    int t= 0;
    int j = 0;
    while(j < d){
        t = t + (X[j]-Z[j])* (X[j]-Z[j]);
        j++;
    }
    double Result = sqrt(t);
    return Result;
}

```

### 2.3 Ramifications of Existing Approach

Existing *OneNN* function calls L2 function which in turn involves square root computation using *sqrt*. Overall N calls to *sqrt* are required to compute N distances. Literature suggests that the square root of an n-digit number can be computed in time  $O(M(n))$  using e.g. Newton's iteration, where  $M(n)$  is the time needed to multiply two n-digit integers. The current best bound on  $M(n)$  is  $n \log n^{2O(\log_* n)}$  provided by Fürer's algorithm [9]. Moreover, existing L2 does not use any short-circuit return statement. This absence of short-circuit termination renders L2 computation time consuming. In order to solve these problems, an innovative algorithm has been designed and implemented.

## 3 Proposed Innovative Approach

In this section, the informal and formal descriptions of our proposed approach is discussed.

### 3.1 Informal Description

As existing algorithm does not allow short circuit functions, we conversely proposed algorithm *OneNN* that is designed in a way that facilitates short circuit terminations. It is important to note that unlike existing algorithm, proposed algorithm attempts to eschew computation of square roots. This algorithm searches the training object Z with smallest L2 squared distance from X. Moreover, eschewing square root allows us to avoid floating point arithmetic.

### 3.2 Formal Description

```

/* Modified OneNN */

const int N = 1000; const int d = 100;
Struct TrainingData{
    int x[N][d];
};

int OneNN(TrainingData T, int N, int* Z){
    int CurrentMinimum = INT_MAX;
    int j = 0;
    int NearestObject = 0;
    while(j < N){
        int t = L2Square((&(T.x[j][0]),Z,d,CurrentMinimum));
        if(t < CurrentMinimum){
            CurrentMinimum = t;
            NearestObject = j;
        }
        J++;
    }
    return NearestObject;
}

int L2Square(int* X, int * Z, int d, CurrentMinimum){
    int t = 0; int j = 0;
    while(j < d){
        t = t + (X[j]-Z[j])*(X[j]-Z[j]);
        if(t >= CurrentMinimum)
            return (INT_MAX); // Short-Circuit
        j++;
    }
    return t;
}

```

## 4 Mathematical Analysis

*OneNN* implements the 1NearestNeighbor algorithm. It will call *L2Square* exactly N times and will avoid exactly N computations of square root operations. How many of these N calls to *L2Square* will allow short circuit? This study suggests that there are three possibilities:

- (i) Best Case: In best case first training object is our nearest object. In this case all calls except the first call will allow short-circuit computation. Thus, (N –

1) calls will allow short-circuit computation. The probability of the best case will be computed as follows:

Total permutations =  $N!$

Permutations where first object is nearest =  $(N - 1)!$

Probability of Best Case =  $(N - 1)! / N! = 1/N$  For large  $N$ , this is low.

- (ii) Average Case: In average case analysis, there is a need to estimate the expected number of times *L2Square* fails to short-circuit the computation. Or equivalently there is a need to know the expected number of times comparison ( $t < \text{CurrentMinimum}$ ) ends up in true. Suppose  $Y$  is the random variable which counts the number of times comparison ( $t < \text{CurrentMinimum}$ ) is true. Also, consider  $Y_j$  is the indicator random variable attached with the case in which above mentioned comparison is true. Then,

$$Y_j = I\{ \text{In } j\text{th iteration } (t < \text{CurrentMinimum}) \text{ is true} \}$$

$$Y_j = 1 \text{ if comparison is true}$$

$$Y_j = 0 \text{ if comparison is false and } Y = Y_1 + Y_2 + \dots + Y_N$$

$$E[Y_j] = \text{probability}\{\text{In } j\text{th iteration}(t < \text{currentMininimum}) \text{ is True}\}$$

Now, there is a need to compute the probability ( $t < \text{CurrentMinimum}$ ) in  $j$ th iteration. Comparison ( $t < \text{CurrentMinimum}$ ) is true when  $j$ th object is nearest 5 among first  $j$  objects. Therefore, we need to calculate the probability of  $j$ th object to be one with the smallest distance from object Z.

Total permutations =  $j!$

Permutations where  $j$ th object is nearest =  $(j - 1)!$

Probability of  $j$ th object being nearest =  $(j - 1)! / j! = 1/j$

$$E[Y_j] = 1/j$$

$$E[Y] = E\left[\sum_{j=1}^N Y_j\right]$$

$$E[Y] = \sum_{j=1}^N E[Y_j]$$

by linearity of expectation

$$E[Y] = \sum_{j=1}^N 1/j$$

$$E[Y] = \log_e(N) + O(1)$$

Therefore, on average  $\log_e(N) + O(1)$  times short circuit will not be possible. Thus, average case analysis suggests that on average  $N - (\log_e(N) + O(1))$  times short circuit will work.

- (iii) Worst Case: In worst case, training objects will have strictly decreasing distance with object Z. In this case, none of the N calls to *L2Square* will allow short circuit. There is a need to estimate the probability of the worst case to occur in a data set. Following calculation help to find that probability estimate of worst case is close to 0 for large N:

$$\text{Worst case permutation} = 1$$

$$\text{Total permutations} = N!$$

$$\text{Probability of Worst case} = 1/N!$$

## 5 Key Results

Our algorithm with innovative approach attempts to avoid  $N \sqrt{}$  calls and allows short-circuit computation. As soon as variable t goes above parameter CurrentMinimum short circuit occurs and as a result a function *L2Square* terminates. On average  $(N - (\log_e N + O(1)))$  calls of *L2Square* will involve short circuit. In worst case, none of the N calls will involve short circuit. However, the probability of such a case is  $1/N!$  (which is close to 0 for large N). In best case,  $(N - 1)$  calls will allow short circuit leading to compute probability of best case as  $1/N$  (which is low for large N).

## 6 Conclusion

Although kNN is an interesting algorithm, it has a fundamental weakness in the form of high computational running time, as there is a need to compute distance with every training object. Furthermore, existing algorithm requires calls to *sqrt*, which are computationally expensive. Moreover, existing algorithm does not exploit the strategy of short circuit computation. Our algorithm attempts to avoid square root computation and floating point arithmetic thereby allowing short circuit-computation for most of the data sets. This study proves the superiority of this approach by analyzing the advantages mathematically. However, as a future work another empirical study of this algorithm is required to measure the gains achieved in running time.

## References

1. M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects. *Science* **349**, 6245 (2015)
2. S. Zhang, et al., Learning k for kNN classification. *ACM Trans. Intell. Syst. Technol. (TIST)* (2017)
3. H. Guo, et al., BPSO-Adaboost-kNN ensemble learning algorithm for multi-class imbalanced data classification. *Eng. Appl. Artif. Intell.* (2016)
4. K. Taunk, S. De, S. Verma, A. Swetapadma, A brief review of nearest neighbor algorithm for learning and classification, in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (2019)
5. B.V. Dasarthy, *Nearest Neighbor Classification Techniques* (IEEE Press, Hoboken, 1990)
6. T.M. Mitchell, *Machine Learning*. McGraw-Hill, Columbus (1997)
7. R.O. Duda, P.G. Hart, D.E. Stork, *Pattern Classification* (Wiley, New York, 2001)
8. T. Hastie, R. Tibshirani, J.H. Friedman, *The Elements of Statistical Learning* (Springer, Berlin, 2001)
9. [https://en.wikipedia.org/wiki/F%C3%BCrer%27s\\_algorithm](https://en.wikipedia.org/wiki/F%C3%BCrer%27s_algorithm)

# Latest Trends in Multi-modality Medical Image Fusion: A Generic Review



**Kapil Joshi, Mohit Kumar, Amrendra Tripathi, Anuj Kumar, Jitender Sehgal, and Archana Barthwal**

**Abstract** Fusion of Medical Images is a simple process to register and merge various images from various modalities of images to enhance the quality of image and reduction in the redundancy for increasing the scalability clinically and capability of images taken for medical purposes to diagnose various medical problems. When we analysis the multi modal medical image fusion algorithm then it increase the diagnosis efficiency and accuracy in clinical order. Multimodal algorithms and systems for medical image fusion show significant achievements in enhancing the accuracy clinically of medical image—based decisions. A factual list of methods is given in this review article and summaries the major challenges faced scientifically in the area of fusion of medical images. A factual list of methods is given in this review article and it gives a summary of the broad challenges faced scientifically in the fusion of areas in medical images. This review also provides the organs details for further the purpose of diagnose system. Research in Fusion of Medical images is defined based (1) on the commonly used methods of image fusion, (2) the modalities of imaging, along with (3) the under-study organ imaging. Conclusively, the paper proposes the latest issues with the working of multi modality fusion of medical images in terms of future perspective.

**Keywords** Multi-modal imaging · Medical imaging · Image fusion · Diagnostics system

---

K. Joshi (✉) · A. Kumar  
Department of CSE, UIT, Uttarakhand University, Dehradun, India  
e-mail: [kapilengg0509@gmail.com](mailto:kapilengg0509@gmail.com)

M. Kumar  
Department of CSE, UIE, Chandigarh University, Mohali, India

A. Tripathi  
Department of CSE, School of Computer Science, UPES, Dehradun, India

J. Sehgal  
G.B. Pant Engineering College Ghurduari, Pauri Garhwal, India

A. Barthwal  
Department of CSE, Women Institute of Technology, Dehradun, India

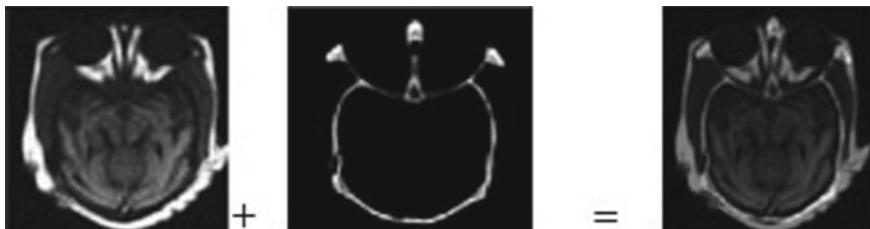
## 1 Introduction

Fusion of Medical Images is a procedure involved in merging of two images that are relevant, from a data set [1] into a fused image. In other terms, Image fusion is a process of producing an image that carries the information stored in many images. The method involves extracting vital details from each individual image and fusing it to create new resulting images. Mainly, the objective of fusion of images is getting the suitable data which becomes understandable by other people. If we discuss about the medical image fusion then Fusion of medical images is the method by which more than one images from one or more than one modalities of images are captured and combined to enhance the consistency of imaging and minimize the random nature and duplication for increasing the application in clinics in terms of medical images to diagnose and evaluation of medical issues.

When we analyze the fusion of medical images, then the merging process is possible with CT, MRI and PET etc. into a single fused resultant image which contains extra essential information from further diagnosis system [2]. Medical image modalities improve the image quality and reduce the clinical medical problem. As there are many modalities capable for the process of medical image fusion. Each image modality provides various information for suitable example is MRIs, X-rays, and ultrasound, CT-Scans, SPECT and PET etc., so as we can see in Fig. 1.

General gap found the limitation of CT, MRI and PET and lack of examination of limited information over there. To overcome the limitation, we may apply the image fusion method.

This paper is organized into five sections, first section is based on introduction and second section gives the details of medical multi-modalities. All human organs information is explained in third section. Comparative study and conclusive part is described in forth and five sections.



**Fig. 1** Medical image fusion (Merging)

## 2 Medical Image Modalities

### 2.1 X-Ray Images

A radiograph or X-Ray is a non-invasive test for diagnosis, which aids the doctors in diagnosis and treatment of various medical conditions. To obtain images of the internal body organs, x-ray imaging includes exposure of a part of the body to an ionizing radiation in small portions.

### 2.2 Ultrasound Images

Ultrasound imaging or Sonography makes use of high-frequency waves of sound to view a suggestive image of the internal body parts. Since the ultrasound images are obtained in real-time, hence they can also show the movement in the internal organs of the body, in addition to the flow of blood through the blood vessels. Like in X-ray imaging, we do not have exposure to ionizing radiation, associated with ultrasounds. Ultrasound technique is a sonar based technique. It uses Ultrasonic sound waves for the diagnosis purpose. This technique is widely used in medical applications due to its safe behaviour and low cost. Its limitation lies in operator reasons such as no air gap should be present in between body [3] and ultrasound probe. Also there should not present any bone structure between the path of the organ under diagnosis area and probe. Ultrasound is mainly used in diagnosis in maternity cases as it has no side effects to the patient and baby.

As each modality image have some different properties. Two or more medical image when fused together will reduce the redundancy and provide accurate diagnosis result. This eliminates the risk of wrong analysis and treatment, enhancing the medical domain achievements.

Ultra Sonography or Ultrasound is a technique based on generating an image through sound waves which have a temporal frequency that is higher. This enables it to produce diagnostic information which is both qualitative and quantitative on the basis of the methodology employed for the specific purpose at hand. The ultrasound technology finds application in various processes like diagnosis of cancer, radiation therapy, tumor diagnosis in the liver, image fusion and biopsy of prostate.

### 2.3 CT Images

A prominent impact has been made in the diagnostic sciences and assessments in the medical field, with the imaging technique called Computerized Tomography. Just like MRI, the CT Scans find use in a varied range of clinical diagnosis. A CT scan comprises of a series of X-Ray images that are taken from multiple angles of the

body and with computerized processing, many images of the cross-sections of the bones, soft tissues and blood vessels are generated. These images provide much more detailed information about the human body than a simple X-Ray might have. The chief advantage of a CT Scan is that it takes much lesser time for scan and have a much higher resolution of the resultant images. The exact levels of radiation in a CT Scan are not understood very well. CT Scans [4] have various other limitations like characterization of limits tissues due to the way an X-Ray probe works, the restrictions that a CT Scan had in traversing the slices needed and the practical drawbacks, like the restricted number of X-Rays that are possible to be developed in short times of scan.

CT Scans have found use in timely diagnosis of the extent of lung damage caused by the Covid 19 in recent times. A simple X-ray may not prove sufficient for the accurate detection of the lung damage state, but a CT scan gives the precise data, which can help the doctors to adjust the treatment for the patient, accordingly.

## 2.4 *MRI Images*

MRI or Magnetic Resonance Imaging has a vital part in diagnosing brain tumours, and is essentially a non-invasive diagnosis. It is a trusted and widely used imaging modality in medical sciences. MRI is a non-invasive type of imaging technique and is efficiently used in diagnosis of the conditions of the brain. These images give one a better look at the pathological soft tissue and the vessel that is a relevant vessel. This finds use in efficiently extracting the abnormal region in the brain, which represents the presence of brain tumour that is reflective.

The biggest advantage of an MRI is that it is a safer method of diagnosis, even for babies and pregnant women. It does not expose the patient to radiation, which can be harmful in the long run. Additionally, there is a much higher accuracy of the images of soft tissues like heart, eyes, brain, etc. An MRI is relatively sensitive to the movement, which makes it a difficult process to assess organs which may be moving, like the heart or the diagnosis of tumours in the mouth.

The image fusion based on MRI is also applied in the studies of prostate. The fusion of MRIs has seen advancements that are methodological, like a similarity in the SSIM (Structure Similarity Match Measure), which helps [5] in improving the accuracy of the application.

There are various other applications, incorporating the fusion based on MRI, including the regeneration of images, liver and lung diagnosis, visualisation of potential field, surgical training and planning, tissue classification, visualisation in multiple dimensions, assessment of breast cancer, extracting colour, structure and shape of the specimen, image registration, pattern visualization and recognition, Gynaecological diagnosis of cancer, Tumour Simulation and MRI guided treatment. The combination that is most prominent is the study of MRI-CT, as the technology is much advanced in this case and is practically usable in the setting of diagnostic clinics.

## 2.5 PET Images

PET Scans or PET Imaging is known as Positron Emission Tomography and is a vital form of imaging in nuclear medicine. Herein, we discuss some of the areas of application where a PET scan is considered as a prime modality in the fusion of data. As in MRI and CT scan, a PET Scan is majorly applied in the studies of radiology in diagnosis of diseases in the brain and their treatment [6].

Using PET, there are a varied range of image fusion application, some being for image segmentation, diagnosis of gynaecological cancer, cervical cancer, simulation of 3D tumor, inertial electrostatic confinement fusion and treatment of cancer, etc.

## 2.6 SPECT Images

SPECT (Single Photon Emission Computed Tomography) Scan is a nuclear method of imaging which is used extensively for the observation of flow of blood to organs and tissues. The areas in which the SPECT Images are applied; neck and head cancer diagnosis, brain treatment and diagnosis, lung cancer treatment, liver diagnosis, detection of tumors, multi-dimensional visualization, bone cancer diagnosis, biopsies, etc. A primary and major challenge in this imaging is the improvement in the image sensitivity with the use of SPECT, in absence any reduction in the resolution of images.

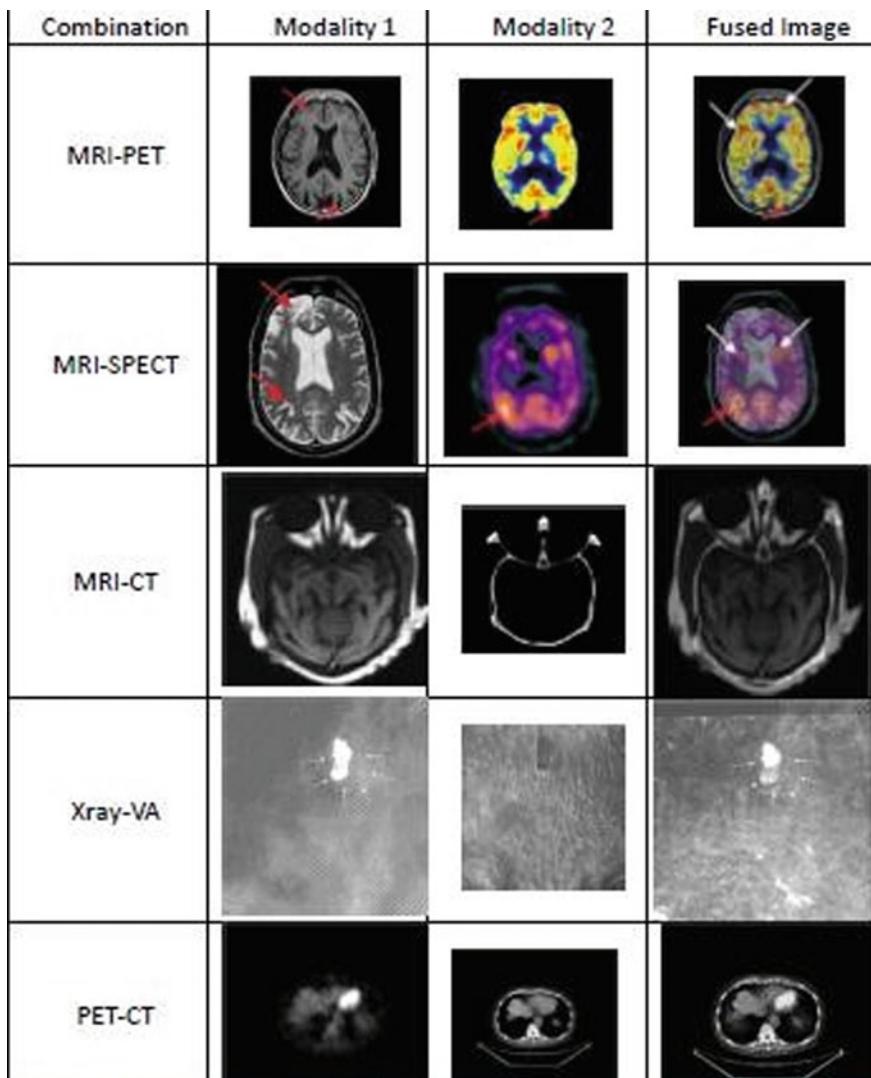
SPECT (Single Photon Emission Computed Tomography) is a very effective tool for nuclear imaging. It can depict the blood flowing to the various organs in the human body. It is used efficiently for brain diagnosis, detection [7] of lung cancer, detection of head and neck cancer, fusion of multimodalities, etc. The image quality in SPECT is poor and is influenced by the noise in image and it needs post processing for enhancing the quality of the image in Fig. 2.

## 3 Application Domain (Organs for Diagnosis)

The various scans previously discussed, find application in the detection of various diseases and the treatment of these diseases in the following organs:

### 3.1 Brain

A wide range of research and analysis has been done on the detection of issues with the brain through medical images [8]. The studies for imaging show various vital pieces of detail about the brain that are not possible to be fathomed by human senses



**Fig. 2** Examples of multi-modality image fusion using modality 1 and modality 2

alone. CT scans is the most commonly used image modality [9] for studying the brain, followed by MRI, DSA, PET, SPECT, MEG, EEG, etc.

In brain studies, fusion of medical images has been used to segment the tissues of brain, brain segmentation based on pseudo coloring in MRIs, Stereotactic brachy therapy of the tumors of the brain, visualization of the cortical potential field, classifying of the different activities of brain, diagnosing Alzheimer's and Parkinson's diseases. Etc.

### **3.2 Prostate**

The use of multi-modal images in medicine is vital in the study of prostate related ailments. There is a wide range of studies and techniques on the image fusion based on prostate, which often face the challenges like the deformation or enlargement of the prostate in setups with multi-modal images. These medical imaging Techniques for analysis of the prostate include the prostate localization for the 3D therapy with conformal radiations, evaluate the volume changes [10] and motion of prostate gland, prostate brachytherapy, prostate tissue histology study, prostate cancer thermal ablation, uroscopy, TRUS, SPECT, PET, fMRI, etc.

## **4 Comparative Analysis**

In the few recent decades, medical imaging has been applied widely to the treatment and clinical therapy as a larger number of image modalities have become available. In general, we divide the medical imaging into functional and structural systems. Structural images like CT scans and MRIs give images in high resolution with information about the anatomy; functional images like SPECT and PER give functional information with a lower resolution spatially. The clinical needs cannot be essentially satisfied with a single image; hence a combination of functional and anatomical images gives more information that is necessary and useful in Table 1.

## **5 Conclusion**

In the past few years, fusion of image fusion in medical imagery has gained much interest as the experts in medical field are guided by the developments in this field. A huge number of modalities in medicine can be an input to the steps of fusion for producing a final image as the output, which is informative. The presented work introduced the multi-modalities medical image fusion with its application and the importance of medical modalities in medical fields. Organs details have been discussed. A literature review of various researches of medical image fusion as well as medical image registration was introduced in this paper.

**Table 1** Comparative analysis among medical image modalities and parameters

	Ultrasound	X-Ray	CT	PET	SPECT	MRI
Investment Cost	Moderate	Moderate	High	High	High	High
Hardware	Moderate	Moderate	High	High	High	High
Complexity						
Resolution	Moderate (0.3–3 mm)	High (~1 mm)	High (~1 mm)	High	High	High (~1 mm)
Frame rate	High (300 fps)	Middle (10–30 fps)	–	–	–	Low (5–10 fps)
Harm to the human body	No	Yes	–	–	–	No
Capability of acquiring a video sequence	Yes	Yes	–	–	–	Yes
Imaged	Mechanical Properties	Tissue absorption	Tissue absorption	–	–	Biochemistry
Penetration	3–25 cm	Through	Through	–	–	Through
Safety	Safe	Ionizing radiation	Ionizing radiation	–	–	Ionizing radiation

## References

- Q. Zhu, H. Li, H. Ye, Z. Zhang, R. Wang, Z. Fan, D. Zhang, Incomplete multi-modal brain image fusion for epilepsy classification. *Inf. Sci.* **582**, 316–333 (2022)
- Y. Xiao, Z. Guo, P. Veelaert, W. Philips, DMDN: Degradation model-based deep network for multi-focus image fusion. *Signal Process. Image Commun.* **101**, 116554 (2022)
- S. Alqahtani, X. Zhang, C. Wei, Y. Zhang, M. Szewczyk-Bieda, J. Wilson et al., Predicting the performance of concurrent systematic random biopsies during image fusion targeted sampling of multi-parametric MRI detected prostate cancer. A prospective study (PRESET Study). *Cancers* **14** (1), 1 (2022)
- G. Makwana, R.N. Yadav, L. Gupta, Comparative analysis of image fusion techniques for medical image enhancement, in *Proceedings of International Conference on Computational Intelligence* (Springer, Singapore, 2022), pp. 241–252
- J. Bhardwaj, A. Nayak, medical image fusion using lifting wavelet and fractional bird swarm optimization, in *Proceedings of the International e-Conference on Intelligent Systems and Signal Processing* (Springer, Singapore, 2022), pp. 277–290
- I. Nazir, I.U. Haq, M.M. Khan, M.B. Qureshi, H. Ullah, S. Butt, Efficient pre-processing and segmentation for lung cancer detection using fused CT images. *Electronics* **11**(1), 34 (2022)
- L. Chandrashekhar, A. Sreedevi, A nature inspired algorithm for enhancement of fused MRI and CT brain images, in *Emerging Research in Computing, Information, Communication and Applications* (Springer, Singapore, 2022), pp. 11–24
- J. Duan, S. Mao, J. Jin, Z. Zhou, L. Chen, C.P. Chen, A novel GA-based optimized approach for regional multimodal medical image fusion with superpixel segmentation. *IEEE Access* **9**, 96353–96366 (2021)

9. K. Wang, M. Zheng, H. Wei, G. Qi, Y. Li, Multi-modality medical image fusion using convolutional neural network and contrast pyramid. *Sensors* **20**(8), 2169 (2020)
10. B. Rajalingam, F. Al-Turjman, R. Santhoshkumar, M. Rajesh, Intelligent multimodal medical image fusion with deep guided filtering. *Multimedia Syst.* 1–15

# Comparison Study of Classification Techniques for Predicting Performance of Students Using Weka Environment



Shivani Sharma and Monika Rathore

**Abstract** Due to the advancement of technology in education systems, significant volumes of data are being stored, making it essential to employ Data Mining methods to enhance the educational process. Data Mining methods help in the classification and prediction of Students' efficiency, absenteeism, and teacher effectiveness. With the assistance of data mining methods, an instructor, as well as learners in course selection and educational administration, can track intellectual development to improve the teaching process. This research compares the three classification algorithms Naive Bayes, J48 decision tree method, and PART algorithm for predicting student's academic performance in the WEKA environment. Knowledge about student performance can be predicted and analyzed by using data mining techniques on student data. This understanding will support the improvement of educational quality, student performance, institute quality, and the reduction of failure rates.

**Keywords** Data mining · Prediction · Decision tree · Classification · Weka tool

## 1 Introduction

Data Mining is a technique for extracting meaningful information from enormous sets of data. To detect and extract patterns from stored data, DM techniques employ a variety of strategies. The pattern discovered will be used to a variety of problems in a variety of sectors, including education, finance, marketing, analytics, healthcare, and games. Because of the vast amount of data stored in those locations, a DM technique is required because the ensuing analysis is significantly more exact and accurate.

---

S. Sharma (✉)  
Rajasthan Technical University, Kota, India  
e-mail: [Shivani2011990@gmail.com](mailto:Shivani2011990@gmail.com)

M. Rathore  
International School of Informatics and Management, Jaipur, India

There has been a surge in interest in using DM to explore the educational profession in recent years. EDM is responsible for developing and evaluating educational content to gain a better knowledge of students' performance. Chronological and administrative data contained in educational establishment databases can be used to compile academic data. Personal or academic data may be collected from students. There are many data mining approaches like Neural Networks, Nave Bayes, K-Nearest Neighbor, Decision Trees that are used in educational data mining. For predicting the accurate student performance, prediction models that encompass all personal, cultural, cognitive, as well as other ambient aspects are required. The ability to accurately anticipate student performance is useful for identifying students with low academic achievement early on. The teacher must provide extra assistance to the identified students to improve their performance in the future. This research paper is going to predict the students' academic performance by examining J48, Naïve Bayes, and PART classification algorithms and will perform comparative analysis of these algorithms using the Weka tool.

## 2 Literature Review

Qasem, Emad, and Mustafa [1] attempted to employ data mining techniques, specifically categorization. They tried to improve the quality of higher education by analyzing information of students, which aided in the research of key attributes that may influence a students' progress in C++ classes. As a methodology, the CRISP framework was used. The performance of three classifiers, Naïve Bayes, ID3, and C4.5 decision tree was compared, and the results revealed that Decision tree C4.5 outperformed the others.

Surjeet and Pal [2] investigated the performance of first-year computer science students using the decision tree algorithms in another study. Each class came up with three forecasts. Students were evaluated on a pass/fail criterion and upgraded based on their performance. This method worked well in determining which students were most likely to fail.

Ahmad et al. [3] created a model for predicting first-year computer science undergraduates' academic results. The information was gathered over eight years, from July 2006–2007 to July 2013–2014. The information gathered covered a wide range of topics, including previous academic abilities, family history, and demography. Decision Tree, Naive Bayes, and Rule-Based Classifiers are used to evaluate student's performance. The trials revealed that the RuleBased Classifier, the most accurate of the other classifiers, with a 71.3% accuracy rate.

Sumitha et al. [4] used a dataset of senior students to construct a model to predict students' academic outcomes. They examined data mining classification techniques and discovered that, based on their data, the J48 algorithm was the best fit for the job.

A study by Khasanah et al. [5] found that strong influence traits can be properly chosen to predict student achievement. Before classifying such a job, feature selection

may be employed. The student statistics came from Universitas Islam Indonesia's Department of Industrial Engineering. They classified and predicted student performance using Bayesian Network and Decision Tree methods. The Feature Selection methods revealed that the first semester's attendance and Grade Point Average were the most important features. When it came to accuracy, the Bayesian Network classification beat out the Decision Tree classification in their case.

To develop classification models, Ankita A Nichat et al. [6] used decision trees and artificial neural networks. To increase production, they used several indicators to determine the students' strengths and shortcomings.

Hilal Almarabeh [7] conducted the other study. He assessed students' academic progress using the WEKA technique and discovered that the volume and type of data affect the accuracy of categorization systems. Bayesian Network, Naïve Bayes, Neural Network, J48, and ID3 classification were used by the author. The Bayesian Network was discovered to outperform others in terms of accuracy.

### 3 Data Mining Classification

Classification is a straightforward procedure for identifying a prototype that distinguishes the key characteristics of data classes or concepts to utilize the model to forecast the class of unknown items. The classification of data is a two-step procedure. The first stage involves analyzing data tuples from basic classifiers using a set of attributes to create a model. The value of the target class attribute is known for each tuple in the training data. To build a model, a classification algorithm is employed for the training data. Test data is used in the second step of classification to ensure that the model is appropriate. The model would be used to categorize the unknown data tuples if its accuracy is adequate. Basic classification approaches include decision trees, Bayesian classification, Bayesian belief networks, and neural networks. This research paper includes a comparative analysis of Naïve Bayes, J48, and PART classification algorithms.

#### 3.1 Naïve Bayes

The Bayes' Theorem is used to generate a list of classification methods known as Naive Bayes classifiers. It's a collection of algorithms that all function on the same premise: every pair of attributes to be classified is distinct from the others. is distinct from the others. The following is a definition of the Bayes theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$P(A)$  is the prior probability, which indicates the likelihood of A occurring, and  $P(B)$  is the marginal probability, which indicates the likelihood of B occurring. The Naive Bayes algorithm is a rudimentary approach for determining possibilities using probabilistic reasoning of each class's features. The fact that the Naive Bayes theorem only takes a modest quantity of training data to estimate the parameters is one of its primary advantages.

### 3.2 J48

J48 is a Java implementation of the standard C4.5 decision tree technique that is open source. J48 uses a predictive machine-learning model-based decision tree to identify the new various attributes from the existing data. A decision tree's internal nodes represent many qualities, while the branches linking the nodes display possible values of these qualities in recorded instances, and the terminal nodes represent the dependent variable's resultant value.

### 3.3 PART

PART is an implied strategy for generating rules. In each iteration, PART creates a pruned decision tree using the C4.5 statistical classifier. The leaves of the best tree are converted into rules.

## 4 Data Set and Attribute Selection

In this study Student data set is being used for experiment analysis in Weka. This data consists of 52 instances and 14 attributes in Table 1. Data is first gathered in an excel sheet. This study is analyzing marks of Second Semester students of the CSE branch in different subjects (Artificial Intelligence, Machine Learning, Data Mining. The target variable is Result and categorized into five values as “Best”, “First”, “Second”, “Third”, “Fail”. The following are descriptions and range values for some of the variables:

Reg\_No: This attribute defines the Registration number of students which is unique for each student.

Gender: The Student's gender. It has two values M-Male, F-Female.

Caste: This attribute presents that a particular student belongs to which category. The value of this attribute is divided into 5 values: Gen—General, SC—Schedule Caste, ST—Schedule Tribe, BC A—Backward Class A, BC B—Backward class B.

**Table 1** Dataset attributes description table

Attribute	Description	Value
Reg_No	Student registration number	Numeric (101–152)
Gen	Gender	M-Male, F-Female
Cst	Caste of student	Gen, SC, ST, BC-A, BC-B, OBC
Loc	Location of house	R-Rural, U-urban
Branch	Branch of student	CSE-Computer Science and Engineering
AI	Marks in Artificial Intelligence	Range 1–100
ML	Machine Learning	Range 1–100
DM	Data Mining	Range 1–100
Percentage	Percentage of student	Percentage of student
Att	Attendance	Poor, average, good
Seminar	Performance in the seminar	Range 1–50
SemP	Performance in the seminar	Best, very good, good, Poor
Result	Result of student	If (percentage $\geq$ 85) then Excellent If (percentage $\geq$ 70 and < 85) then very good If (percentage $\geq$ 55 and < 70) then good If (percentage $\geq$ 35 and < 55) then Pass If (percentage < 35) then fail

Location: Location of the student. It has two values: Rural and Urban.

Branch—CSE.

Total\_Marks: Total Marks obtained in three different subjects.

Percentage: Percentage obtained by dividing obtain marks by total marks and multiplied by 100.

Result: It is categorized in Five values. Best  $\leq$  85, First  $\geq$  70 and < 85, Second  $\geq$  55 and < 70, third  $\geq$  40 and < 55, Fail < 40.

SemP: Performance in seminar and it is evaluated through three class values: Best, very good, good, Poor.

Att: student's attendance

## 5 Result and Discussions

This paper includes a dataset of 52 students to compare three classifiers: Naive Bayes, j48, and PART. This examination is carried out with the help of the Weka machine learning framework, which comprises a set of popular learning schemes that can be applied to real-world data mining. Experiments are run on Weka and



**Fig. 1** WEKA screenshot of data distribution in the preprocessing stage

cross-validated ten times. For investigating the classifier's accuracy, tenfold cross-validation is statistically acceptable. The first approach is to determine the instances in number (correctly and incorrectly) of the student dataset by using Naïve Bayes, J48, and PART classification algorithm and then we will evaluate the classification accuracy and confusion matrix. The information regarding the real and expected classification is described in the form of a confusion matrix. In this section, a comparison of classification accuracy results of the Naïve Bayes, J48, and PART techniques has been performed. By applying Naive Bayes, J48 and, PART algorithm on Student's dataset, the obtained results are shown as below in Figs. 2, 3, and 4, respectively. Table 2 compared the performance parameters of these three algorithms.

## 6 Conclusion and Recommendation

Data mining will be particularly useful in the context of education. The difficulty of predicting a student's academic achievement is a major challenge for academic institutions. While predicting student performance, data mining approaches and methodologies are useful for detecting students' skills, interests, and limits. This study reveals that J48 consumes lesser time to develop the model, has more correctly categorized instances and has a higher predictive performance than other two. Consequently, it can be inferred that j48 outperformed on the dataset provided by the

```

Classifier output

==== Summary ====
Correctly Classified Instances      47          90.3846 %
Incorrectly Classified Instances    5           9.6154 %
Kappa statistic                   0.8547
Mean absolute error               0.056
Root mean squared error          0.2021
Relative absolute error           16.5414 %
Root relative squared error     49.2008 %
Total Number of Instances        52

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|
| 0.500         | 0.021   | 0.667   | 0.500     | 0.571  | 0.548     | 0.823 | 0.571    | Best     |       |
| 0.933         | 0.081   | 0.824   | 0.933     | 0.875  | 0.823     | 0.948 | 0.789    | First    |       |
| 0.889         | 0.000   | 1.000   | 0.889     | 0.941  | 0.932     | 1.000 | 1.000    | Third    |       |
| 0.958         | 0.036   | 0.958   | 0.958     | 0.958  | 0.923     | 0.997 | 0.997    | Second   |       |
| Weighted Avg. | 0.904   | 0.041   | 0.904     | 0.904  | 0.902     | 0.867 | 0.970    | 0.905    |       |


==== Confusion Matrix ====


|   |    |   |    |                   |
|---|----|---|----|-------------------|
| a | b  | c | d  | <-- classified as |
| 2 | 2  | 0 | 0  | a = Best          |
| 1 | 14 | 0 | 0  | b = First         |
| 0 | 0  | 8 | 1  | c = Third         |
| 0 | 1  | 0 | 23 | d = Second        |


```

**Fig. 2** The outcome of the Naïve Bayes classification algorithm

```

Classifier output

==== Summary ====
Correctly Classified Instances      48          92.3077 %
Incorrectly Classified Instances    4           7.6923 %
Kappa statistic                   0.8861
Mean absolute error               0.0385
Root mean squared error          0.1961
Relative absolute error           11.3571 %
Root relative squared error     47.7509 %
Total Number of Instances        52

==== Detailed Accuracy By Class ====


|               | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC   | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-------|----------|----------|-------|
| 1.000         | 0.021   | 0.800   | 1.000     | 0.889  | 0.885     | 0.990 | 0.800    | Best     |       |
| 0.933         | 0.027   | 0.933   | 0.933     | 0.933  | 0.906     | 0.953 | 0.890    | First    |       |
| 0.889         | 0.023   | 0.889   | 0.889     | 0.889  | 0.866     | 0.933 | 0.809    | Third    |       |
| 0.917         | 0.036   | 0.957   | 0.917     | 0.936  | 0.884     | 0.940 | 0.915    | Second   |       |
| Weighted Avg. | 0.923   | 0.030   | 0.926     | 0.923  | 0.924     | 0.887 | 0.947    | 0.881    |       |


==== Confusion Matrix ====


|   |    |   |    |                   |
|---|----|---|----|-------------------|
| a | b  | c | d  | <-- classified as |
| 4 | 0  | 0 | 0  | a = Best          |
| 1 | 14 | 0 | 0  | b = First         |
| 0 | 0  | 8 | 1  | c = Third         |
| 0 | 1  | 1 | 22 | d = Second        |


```

**Fig. 3** The outcome of the J48 classification algorithm

```

Classifier output
==== Summary ====
Correctly Classified Instances      46          88.4615 %
Incorrectly Classified Instances   6           11.5385 %
Kappa statistic                   0.8277
Mean absolute error               0.0577
Root mean squared error           0.2402
Relative absolute error            17.0356 %
Root relative squared error       58.4827 %
Total Number of Instances         52

==== Detailed Accuracy By Class ====
          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC    ROC Area  PRC Area  Class
          0.750    0.042    0.600     0.750    0.667     0.640    0.854    0.469    Best
          0.933    0.027    0.933     0.933    0.933     0.906    0.953    0.890    First
          0.778    0.023    0.875     0.778    0.824     0.791    0.877    0.719    Third
          0.917    0.071    0.917     0.917    0.917     0.845    0.923    0.879    Second
Weighted Avg.      0.885    0.048    0.890     0.885    0.886     0.838    0.918    0.823

==== Confusion Matrix ====
      a  b  c  d  <-- classified as
 3  0  0  1 |  a = Best
 1 14  0  0 |  b = First
 1  0  7  1 |  c = Third
 0  1  1 22 |  d = Second

```

**Fig. 4** The outcome of the PART algorithm

**Table 2** Parameters of performance and their values for Naïve Bayes, J48 and PART

Parameters	Naïve Bayes	J48	PART
Time taken to build a model (in seconds)	0	0.01	0.03
Correctly classified instance	47	48	46
Incorrectly classified instance	5	4	6
Kappa statistics	0.8547	0.8861	0.8277
Mean absolute error	0.056	0.0385	0.0577
Root mean square error	0.2021	0.1961	0.2402
Relative absolute error (%)	16.54	11.35	17.03
Root relative squared error (%)	49.20	47.75	58.48
Accuracy	90.38	92.30	88.46
Total number of instances	52	52	52

students. Aside from the students' dataset, it can also be deduced that several methods of classification are optimized for specific dataset categories.

## References

1. Q. Al-Radaideh, M. Al-Shawakfa Emad, M. Al-Najjar, *Mining Student Data Using Decision Trees*. International Arab Conference on Information Technology (2006)
2. Y.S. Kumar, P. Saurabh, Data mining: prediction for performance improvement of engineering students using classification. *World Comp. Sci. Info. Tech. J.* **2**, 51–56 (2012)
3. S. Mushtaq, S.N. Singh, in *Implementation of Classification Techniques in Education Sector for Prediction and Profiling Students*. Emerging Technologies in Data Mining and Information Security, pp 435–444
4. A. Peña-Ayala, *Educational Data Mining: Applications and Trends*, vol. 524 (Springer, New York, 2013)
5. K. Balwinder, G. Anu, R.K. Singla, Applicability of traditional classification techniques on educational data. *IJRTE* **8**(6), 2277–3878 (March 2020)
6. W. Fairos, W. Yaacob, Supervised data mining approach for predicting student performance. *Indonesian J. Electrical Eng. Comp. Sci.* **16**(3), 1584–1592 (2019)
7. L. Guarin., C. Ernesto., E.L. Guzman., F.A. Gonzalez, A model to predict low academic performance at a specific enrollment using data mining technologies, del Aprendizaje. *IEEE RevistaIberoamericana de* **10**(3), 119–125 (2015)

# Multimodal Autism Detection in Children Through Facial and Speech Emotion



Monalin Pal and P. Rubini

**Abstract** Autism Spectrum Disorder is a neurological disorder which affects the emotional, cognitive, social and physical health of the individual. Symptoms involve trouble in communication, difficulty in social interactions, obsession and repetitive behaviors. Identifying ASD at early age is important to reduce the effects of disorder through behavioral, educational and family therapies. Artificial Intelligence has played a significant role in identifying ASD in children. This research presents a new approach to identify ASD through emotion in children through video and audio using Artificial Neural Network. Model is trained using Convolutional Neural Network to detect facial emotion on images of autistic and non-autistic children and further use to generate emotion features. Speech Emotion Recognition is trained to predict the emotion of the child and get speech and emotion features. Face and speech emotion features are trained using supervised algorithms to predict ASD. Logistic Regression outperforms the other model with an F1 Score of 0.70 for Non ASD and 0.60 for ASD.

**Keywords** Autism detection · Facial emotion recognition · Speech emotion recognition · Neural networks

## 1 Introduction

Autism Spectrum Disorder is a type of condition which affects the social communication and interaction in the children and adults [1]. Other characteristics include atypical behavioral activities like difficulty in transition from one activity to another, unusual reactions to sensations and difficulty in focusing on tasks. Autism has adverse effects on education and employment opportunities. Autism can be spotted early in

---

M. Pal (✉) · P. Rubini

Department of Computer Science and Engineering, CMR University (CMRU), Bangalore,  
Karnataka 560043, India

e-mail: [monalinpalr@gmail.com](mailto:monalinpalr@gmail.com)

P. Rubini

e-mail: [rubini.p@cmr.edu.in](mailto:rubini.p@cmr.edu.in)

childhood but it is often not treated until much later. One in 270 people has Autism globally and some studies has reported figures which are much higher. ASD in much of the low- and middle-income countries is unknown. ASD is caused by many factors which can be environmental and genetic. ASD detection and interventions from early childhood can improve the health, development and mental well-being of the people.

Timely access to therapy and interventions can improve the ability to communicate effectively and interact socially.

Expenses for autistic children is high and complex which includes health care, educational expenses, counselling, therapy etc. ASD in children can been seen in all countries including low- and middle-income countries. In low- and middle-income countries, access to counsellors and therapists is very minimal and which may aggravate the effects of disorder in children if not recognized early. This paper discusses a novel multimodal framework to identify ASD through video and audio using facial emotion features and speech emotion features. By this research, it would be helpful for the parents and caretakers to assess their children for ASD who are located remotely using video and audio, and take necessary steps for the therapy.

## 2 Literature Review

Recently artificial intelligence has addressed problems in various fields like image classification, speech recognition, text classification and other areas. Many researches have been done for autism detection using machine learning and deep learning techniques. Less researches have been done for autism detection through multimodal approach by using emotion only. Autism detection is done on various types of data sources like image, speech, video, brain MRI scans, genetic data, behavioral data and through text with contextual words. This section reviews about the studies that is done for autism detection through facial image and speech emotion recognition.

### 2.1 Facial Emotion Recognition

Face processing for children with ASD is a tough task. Ability of children with ASD to recognize facial expressions is weakened which causes difficulties in social interaction [2]. Studies [3] and [3] classify and understand the facial expressions when compared with regular children of the same age. Study uses static front view in the images. Children with ASD share mutual facial attributes like wide upper face, broad-set eyes, shorter center area of face, cheeks and nose when compared with the normal children. Reference [4] classifies facial images in to four categories using back-propagation neural networks and was able to achieve 90.52% accuracy for training and 81.58% for the test images. Local Binary Patterns (LBP) were used in [5] to describe faces through SVM and they achieved an accuracy of 94.81% to identify the gender. Reference [6] proposes an architecture with CNN and Mobilenet

backbone architecture which achieved an accuracy of 94.6%. Authors in [7] propose ASD detection by combining eye gaze movements and facial emotion recognition. They propose a supervised learning method to detect ASD created on features for gaze information, task details and facial features using neural network and attained an accuracy of 86%.

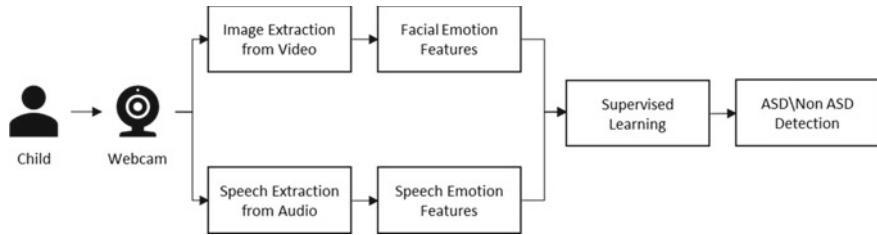
## 2.2 *Speech Emotion Recognition*

Speech emotion recognition forms a vital part of human information processing.

Emotion recognition from speech has become popular lately and significant researches have been done to identify the emotion by improving the accuracy. Reference [8] classified 8 types of emotion namely calm, angry, sad, happy, fearful, disgust and surprised by extracting F0 counter, spectral envelope, and aperiodic information from the data. In another study [9], Linear Discriminant Classifiers and K Nearest Neighbor were used and fear and neutral gave highest accuracy up to 92.5% and 97.1% respectively using LDC model. Authors in [10] re-sampled the data at a frame rate of 40 ms and suggested a Neural Network model by combining Convolutional Neural Network and Long Short-Term Memory networks. Precision of Arousal was 78.7% and valence was 44%. Hamming window function was used to generate spectrograms using Short Time Fourier Transform [11]. Study applied Attention-BLSTM-FCN model and achieved an accuracy of 68% on the dataset. Reference [12] proposes a system to identify emotion using CNN with data augmentation and accuracy for train, test and validation is 0.857, 0.743 and 0.756 respectively. Authors in [13] propose an architecture linking both convolutional layers, high-level features from spectrograms, and recurrent ones for combining long-term dependencies. Techniques of data augmentation with layerwise optimizer adjustment, vocal track length perturbation and batch normalization is performed to achieve a weighted accuracy of 64.5% and an unweighted accuracy of 61.7% for four emotions. Authors in [14] propose a classification model for emotion recognition from speech of autistic children for Tamil Language. They used MFCC features and trained using SVM.

## 3 Proposed Methodology

We propose a novel framework to identify the autism in children through video and audio modalities using facial and speech emotion features. Architecture for autism detection is shown in Fig. 1. Real-time video and audio are captured through a webcam that will be given as input to the framework. Images are extracted from the frames of the video and sent to the facial emotion recognition module to get facial emotion features. Similarly, speech data is extracted from audio and fed to the speech emotion recognition module to get speech emotion features. Features from both the modules are trained using supervised learning techniques for classification.



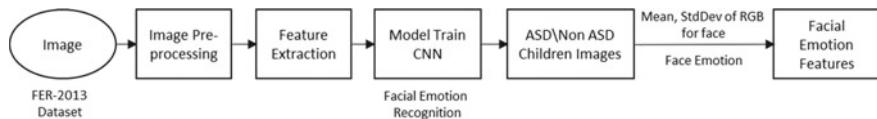
**Fig. 1** Architecture for Autism detection

### 3.1 Facial Emotion Features Extraction

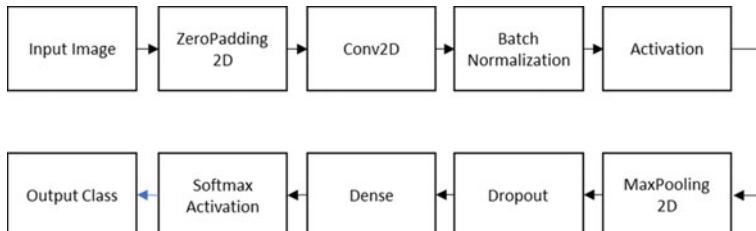
**Data Collection:** Autism dataset consist of facial images for autistic and non-autistic children from Kaggle competition provided by Gerry Piosenka [15]. Dataset has 1468 jpeg images for both autistic and non-autistic children. For facial emotion recognition model training, we used dataset from FER-2013 dataset [16].

**Model Architecture:** Pipeline is shown in Fig. 2 which has components for image preprocessing, feature extraction, model train using CNN for facial emotion recognition and predicting on dataset for autistic and non-autistic children's images to get facial emotion features. Image preprocessing involves correcting the original image by removing noise, improving the exposure and brightness by using OpenCV [17] followed by feature extraction.

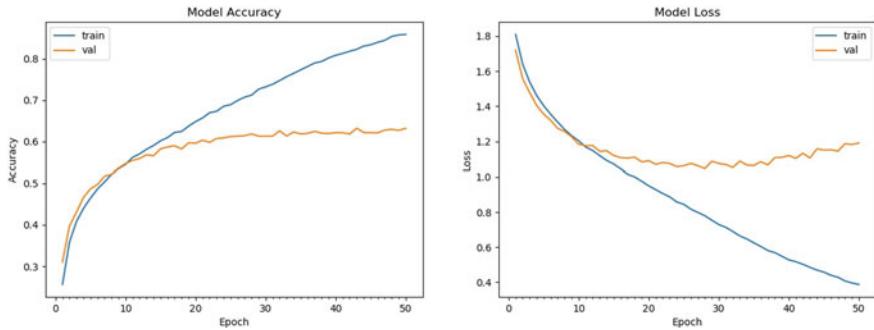
**Model Train:** Model architecture is shown in Fig. 3 for Facial Emotion Recognition for 7 classes namely angry, disgusted, fearful, happy, neutral, sad and surprised. Two Fully connected convolutional layers are appended at end of the model. Dense layer



**Fig. 2** Pipeline for facial features extraction



**Fig. 3** CNN model architecture for facial emotion recognition



**Fig. 4** Model accuracy and loss for facial emotion recognition

**Fig. 5** Facial emotion recognition



with 128 neurons using  $L2 = 0.015$  and ReLu activation function is then connected to prediction layer by using Softmax activation function.

Model accuracy and loss is shown in Fig. 4. 50 epochs were used during training. Overall accuracy of the model on the training dataset is 0.89 and for test dataset is 0.62 for facial emotion recognition.

Facial Emotion Recognition on Autistic children's images is shown in Fig. 5. Facial image features are extracted within the boundary boxes of the individual faces by calculating the standard deviation and mean for the images and combined with facial emotion to get facial emotion features that will be fed to the supervised model.

### 3.2 Speech Emotion Features Extraction

**Data Collection:** Dataset used is Berlin Emotional Speech Database (Emo-DB) [18] for training the speech emotion recognition model and it consists of a database of emotional utterances spoken by actors. It has emotions for anger, happiness, sadness, fear, disgust and neutrality. Speech dataset for Autistic children is used from [19] which has the recorded voices in mp3 format.

**Model Architecture:** Architecture is shown in Fig. 6 which has components for MFCC feature extraction, CNN model for Speech Emotion Recognition to predict on autistic\`non autistic speech data to get speech emotion followed by speech features extraction for the supervised model.

**Model Train:** Architecture involving both convolutional layers with MFCC features extracted from speech is used for training the model and is derived from [13]. Model architecture is shown in Fig. 7.

Mel Frequency Cepstral Coefficients (MFCC) features are generated from the input speech and model architecture has 1–6 convolutional layers, Bi-LSTM with 1–4 layers and finally softmax activation function with a dense layer for classification. For optimization, stochastic gradient descent (SGD) and a batch size of 16 is used along with L2 regularization. Overall model accuracy is 86% with 50 epochs. Spectrogram for sad emotion, model loss and model accuracy are shown in Fig. 8.

Speech features for supervised model is derived from input speech of autistic\`non autistic data which includes statistical features like mean, median, standard deviation etc. and energy, zero-crossing rate and tempo of the speech. These features along with speech emotion are fed to the supervised model for ASD/non ASD prediction.

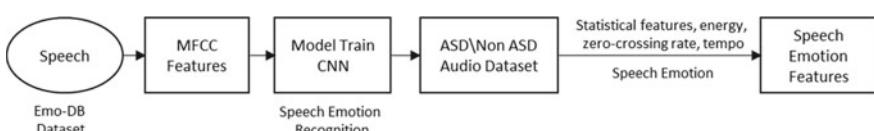
**Energy:** Total magnitude of the signal

$$E(x) = \sum_n |x(n)|^2 \quad (1)$$

**Zero-Crossing Rate:** No of times a signal crosses horizontal axis. Below formula is using librosa library in Python.

$$\text{zero\_crossings} = \text{sum}(\text{librosa.zero\_crossings}(x, \text{pad=False})) \quad (2)$$

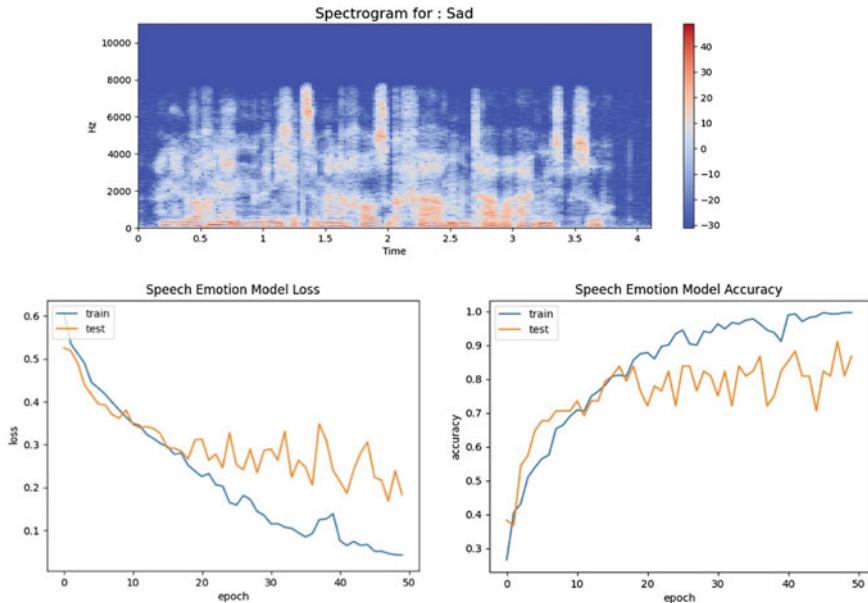
**Tempo:** Estimate of the tempo in beats per minute.



**Fig. 6** Speech emotion recognition architecture



**Fig. 7** Model architecture for emotion recognition from speech



**Fig. 8** Sample spectrogram, model loss and model accuracy

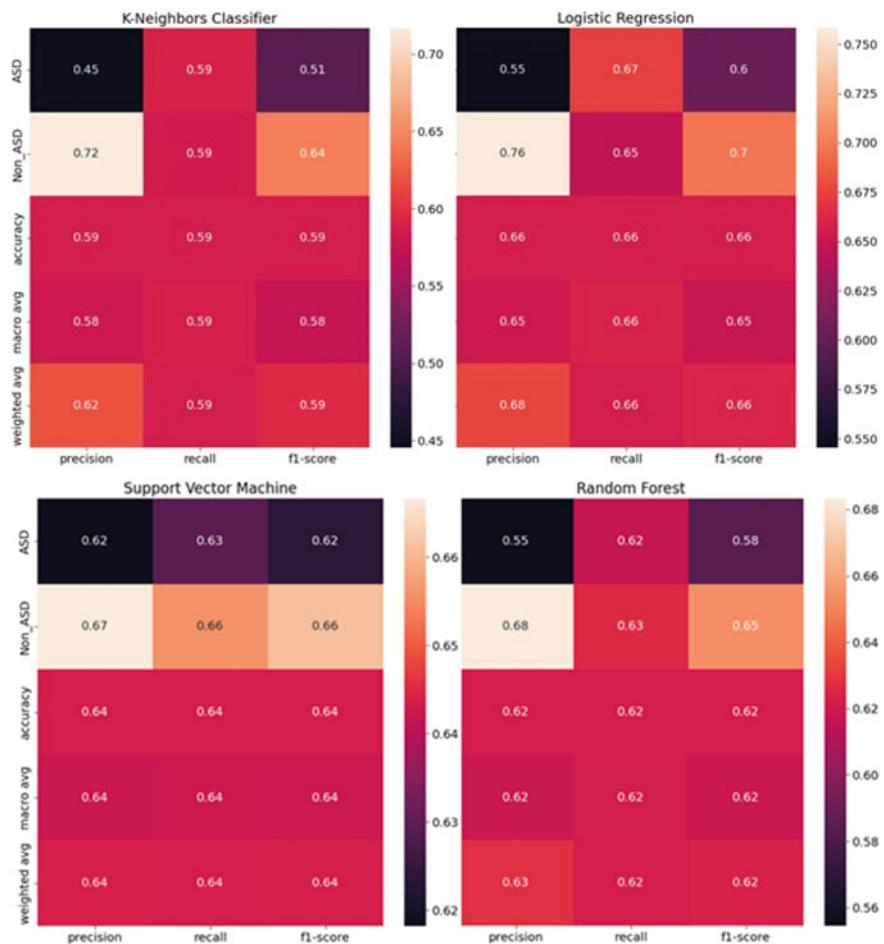
```
tempo = librosa.beat.tempo(x)[0] (3)
```

### 3.3 Supervised Models for ASD Detection

Facial emotion features and speech emotion features from the previous steps were used to build a binary classification model using supervised techniques like Logistic Regression, Random Forest, K-Neighbors Classifier and Support Vector Machine. We had total of 22 features for both facial and speech features combined. Classification reports for all the methods are shown in Fig. 9. We could see that Logistic Regression outperforms the other techniques with an F1 score of 0.70 for Non ASD and 0.6 for ASD.

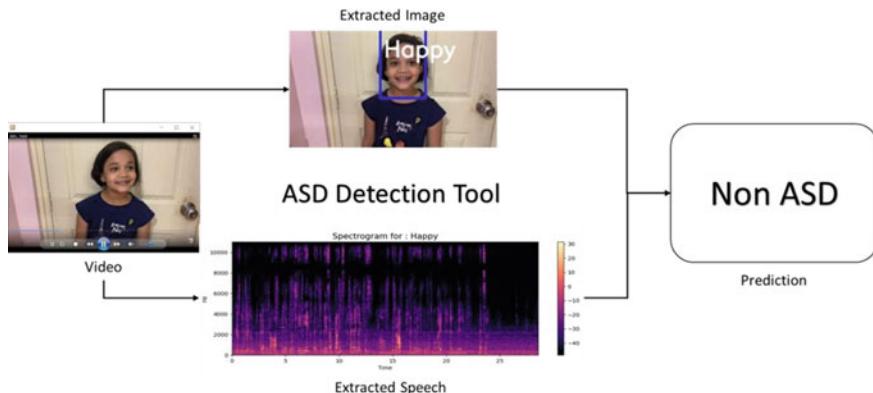
## 4 Results and Discussion

Child is asked to sit in front of the computer and we capture the video and audio from webcam and mic for 30 s duration. Once its captured, video is split in to image frames for facial emotion recognition followed by facial features extraction and audio is fed



**Fig. 9** Classification report for supervised techniques

to the speech emotion recognition to get speech features. Once both the features are available, it will be fed to the supervised model to get the ASD\Non ASD Prediction. Realtime framework which helps to detect the ASD\Non ASD is shown in Fig. 10. Real time pipeline is validated with 10 children of whom 5 were with ASD and another 5 with Non ASD. Videos of 30 s duration were recorded through webcam and saved on the disk and run through the pipeline to detect ASD\Non ASD. We were able to achieve an F1 Score of 0.71 for Non ASD and F1 Score of 0.62 for ASD which outperforms the previous multimodal techniques to detect Autism.



**Fig. 10** End to end real-time ASD detection

## 5 Conclusion

This research presents a new approach to identify ASD in children through real-time using video and audio. Current study discussed about techniques implemented to identify the autism through facial emotion in children and also discussed methodology about identifying emotion from the speech in real time. This research would be helpful to the parents/caretakers who are devoid of health care and far reach to counsellors. Person with a smart phone or through computer can use the tool remotely and self-assess for identification of ASD. Study achieved an accuracy of 0.62 for facial emotion recognition and 0.86 for speech emotion recognition. For Autism detection, Logistic Regression outperformed the other models and achieved an F1 Score of 0.70 for Non ASD and 0.60 for ASD. Limitations include lack of video and audio data for training the model. Future scope involves adding more modalities like pose estimation, eye gaze and point feature analysis on facial features to improve the accuracy of ASD detection.

## References

1. WHO Homepage. <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>. Last accessed 03 Mar 2021
2. B. Gepnerr, C. Deruelle, S. Grynfeltt, *Motion and Emotion: A Novel Approach to the Study of Face Processing by Young Autistic Children*, p. 11
3. S.J. Weeks, R.P. Hobson, The salience of facial expression for Autistic children. *J. Child Psychol. Psychiatry* **28** (1), 137–152
4. R.P. Hobson, The Autistic Child's appraisal of expressions of emotion: a further study. *J. Child Psychol. Psychiatry* **27** (5), 671–680
5. W.B. Horngg, C.P. Le, C.W. Chenn, Classification of age groups based on facial features. *淡江理工學刊* **4** (3), 183–192 (2001)

6. C. Shan, Learning local binary patterns for gender classification on real-world face images. *Pattern Recogn. Lett.* **33**(4), 431–437 (2012)
7. F. Burkhardt, A. Paeschke, M. Rolfes, W. Sendlmeier, B. Weis, A database of German emotional speech, in *Proceedings Interspeech 2005*, Lissabon, Portugal (2005)
8. M. Beary et al., Diagnosis of Autism in children using facial analysis and deep learning. arXiv: Computer Vision and Pattern Recognition (2020): n. pag.
9. R. Aloufi, H. Hadadi, D. Boylee, Emotionless: Privacy-Preserving Speech Analysis for Voice Assistants. arXiv preprint [arXiv:1908.03632](https://arxiv.org/abs/1908.03632) (2019)
10. M. Bojanić, V. Delić, A. Karpov, Call redistribution for a call center based on speech emotion recognition. *Appl. Sci.* **10** (13), 4653 (2020)
11. P. Tzirakiss, J. Zhang, B.W. Schuller, End-to-end speech emotion recognition using deep neural networks, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2018), pp. 5089–5093
12. Z. Zhao, Z. Bao, Y. Zhao, Z. Zhang, N. Cumins, Z. Ren, B. Schuller, Exploring deep spectrum representations via attention-based recurrent and convolutional neural networks for speech emotion recognition. *IEEE Access* **7**, 97515–97525 (2019)
13. S.N. Zisad, M.S. Hossain, K. Andersson, *Speech Emotion Recognition in Neurological Disorders Using Convolutional Neural Network*. [https://doi.org/10.1007/978-3-030-592776\\_26](https://doi.org/10.1007/978-3-030-592776_26) (2020)
14. M. Jiangg, S.M. Franciss, D. Srishyla, C. Conelea, Q. Zhao, S. Jacob, Classifying individuals with ASD through facial emotion recognition and eye-tracking, in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2019), pp. 6063–6068. <https://doi.org/10.1109/EMBC.2019.8857005>
15. G. Bradski, The OpenCV Library. Dr. Dobb's Tools (2000)
16. P. Hendrikks, C. Koster, J.C.J. Hoekss, Referential choice across the lifespan: Why children and elderly adults produce ambiguous pronouns. *Language Cogn. Neurosci.* **29**(4), 391–407 (2014)
17. C. Etienne et al. *CNN+LSTM Architecture for Speech Emotion Recognition with Data Augmentation* (2018)
18. Gerry, *Detect Autism from a Facial Image*. Retrieved 03 May 2021 from <https://www.kaggle.com/gpiosenka/autistic-children-data-set-traintestvalidate>
19. C. Sunitha, R. Ponnusamy, *Recognising and Classify Emotion from the Speech of Autism Spectrum Disorder Children for Tamil Language Using Support Vector Machine*, pp. 25587–25602 (2014)

# A Contemporary Learning on Smart Devices for Smart Metropolis



Sonali Vyas, Shaurya Gupta, Sourabh Singh Verma,  
and Vinod Kumar Shukla

**Abstract** Home Automation is used to remotely control household appliances to minimize efforts. A home security system has the advantage of protecting your home from fire and diversion. A smart wireless home system plays an important role in a person's life and enhances popularity due to its flexibility, versatility and low cost of installation. Although there are a variety of technologies used in all mobile technology, it plays an important role, automated household appliances or awareness in dangerous situations. The attractive GUI can be used for a smart home system, accessible with smartphones, tablets, laptops and PC. This study will provide updated knowledge about the latest and upcoming technologies and their usability's in the technical society we live in. Perhaps the most important re-source of the city for its citizens. However, in a few cases, smart city plans are initiated without citizen analysis of the improvements made in their city. We discuss varied existing and proposed smart devices application in smart city framework implementation for enhancing citizen contribution for supporting the development of the Smart City plan. To provide hypothetical foundations and strong support for the planning and development of future cities.

**Keywords** Smart metropolis · Smart living · Smart household · Smart devices · Smart cities · Smart infrastructures · Smart technologies

---

S. Vyas · S. Gupta (✉)

School of Computer Science, University of Petroleum and Energy Studies, Dehradun, India  
e-mail: [shaurya.gupta@ddn.upes.ac.in](mailto:shaurya.gupta@ddn.upes.ac.in)

S. S. Verma  
SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India

V. K. Shukla  
Amity University, Dubai, UAE

## 1 Introduction

Today's world has been running with such a speed that people have to really manage their resources and specially time to perform even most basic and simple tasks of their daily life. This management is mostly seen in smart cities because of the work load and the lack of time people are facing in order to be more productive. A smart home plan is very beneficial in everyday life as it reduces a person's work, saves electricity and reduces home safety concerns for working people. Focus on automation to control light mode ON/OFF, fan speed and other household items remotely. Home security includes services such as gas leaks and ground protection. This program is very helpful for the elderly and the disabled and the working people, it is a blessing as it alerts a person to a bad situation that has arisen at home in his absence. That is why smart devices are there to help people of the smart cities live a better and secured life [1]. Some examples of these devices may include smart lighting system in homes, amazon echo, and mobile smart devices like personal digital assistants (PDAs), portable game consoles with wireless connectivity and there are a lot more to list. As per ongoing urban migration patterns, cities are being defied to offer a decent standard of living, as well as prioritize investment in urban infrastructure, which requires efficient management. Though, with the growth of the concept of smart cities, Information Technologies and multi-sectoral approaches are helping cities improve their services and use of infrastructure, looking to maintain balance in terms of social, economic and environmental impact. The process for cities to work to transform into smart cities must consider technology and sustainability to be an integral part of their lives. In terms of technology, cities should use information technology, as well as more sophisticated analytical tools, to develop a citizen-centered service approach where the citizen should be at the center of change, becoming the main beneficiary of the city's new paradigm. Therefore, citizens should be consulted and through citizen satisfaction assessments, they can understand the needs of citizens and help develop urban-focused cities.

## 2 Smart Metropolis

According to English dictionary, a smart is one who is having or showing a quick-witted intelligence. A smart city is an urban area that uses information and communication technology to share information with people and help the government to improve the quality of operational efficiency for citizen welfare [2]. Smart city technology is increasingly being used to improve public safety, from monitoring the areas of high crime rate to improving emergency preparedness with sensors. Smart buildings are also a part of a smart city project. Sensors, which are installed in buildings and other structures to detect wear and tear and notify officials when repairs are needed. VANET and MANET are some of the technologies, which helps in making smart cities better [3]. Now, what makes smart cities more reliable? Because it is not

possible to rely on just the government technologies to live a better lifestyle. The smart device makes our lives much easier, saves time and effort.

### 3 Smart City Oriented Citizens

The city should recognize the requirements of its citizens in an urban area. Else, cities drop into un-successful infrastructure development processes and define advanced public policies. As technology speeds up city resources, citizens are able to communicate their ideas on social media, with some activities proposing to have ontologies for managing all that citizen communication to ensure the validity of public policies [4]. Though, for achieving the optimal use of technology to establish connection between the city with its inhabitants, it is essential to have a procedure to measure experience of user for city facilities and their expectations. In order to describe and apply such metrics, it is essential to recognize maturity models for cities in relation to its inhabitants like: dispersed, integrated and connected [5]. The distributed model is where the government is the key decision-maker in the city's infrastructure services. In this model, the processes established to communicate with citizens and receive their response do not work. It could be due to the digital diversity of a significant number of people, as well as the way society is organized. When ICT infrastructure is installed, the management model is put in one place to be effective in solving citizen problems. The second level of maturity is well-defined as unified, which results when the ICT infrastructure is deployed, digital diversity in the community is minimum and there are precise procedures for collecting citizens' views on all significant funds in city infrastructure or services. People are highly planned, establishing connections with the government via social media platforms. Lastly, the maximum maturity level is the connected model, between government and citizens. In this, open practice is followed by government in which public reviews decisions about city infrastructure (Table 1).

### 4 Smart City and Artificial Intelligence

- **Smart House:** A smart house or a smart home is a house equipped with smart appliances like lightning, air conditioning, smart furniture and entertainment resources like TVs, audio and video systems, voice control and sensors for temperature control and many more. Originally, smart home technology was used to control environmental systems such as lighting and heating, but recently the use of technology has developed so much that almost any electrical equipment within the house can be included in the system [6].
- **Smart Living:** Smart living means living and leading a better approach of life. It is the way of living smartly in a smart house built in a smart city surrounded by smart devices [7]. If it called 'smart' living, it means most of the physical objects

**Table 1** Contrast amongst smart metropolises metrics models [3]

Smart metropolis factors	ISO indicators for smart metro quality of life
Smart economy	<ul style="list-style-type: none"> <li>• Economy</li> <li>• Finance</li> </ul>
Smart environs	<ul style="list-style-type: none"> <li>• Energy</li> <li>• Environs</li> <li>• Water Management</li> <li>• Solid waste management</li> </ul>
Smart existence	<ul style="list-style-type: none"> <li>• Fitness</li> <li>• Wellbeing</li> <li>• Housing</li> <li>• Emergency Response</li> </ul>
Smart kinesis	<ul style="list-style-type: none"> <li>• Transference</li> <li>• Urban Planning</li> </ul>
Smart society	<ul style="list-style-type: none"> <li>• Edification</li> <li>• Recreation</li> </ul>
Smart administration	<ul style="list-style-type: none"> <li>• Governance</li> </ul>

around us are connected to the web in one way or another, development that is radically changing the way we live. From ‘robots’ that chop off the vegetables, to detectors that check how long one sleeps, smart devices are becoming more and more a part of daily life. Smart living [8] is about taking in consideration the following concepts:

- **Smart water:** It make decisions that are more informed to the administrators about protecting city’s water supply and prevent water waste using data and detecting water pressure, temperature etc.
- **Smart building:** It involves optimization in terms of construction power usage with the help of motion sensor lights, which can dim or shut off as and when room is empty.
- **Smart environment:** It involves monitoring air quality, pollution, and other weather conditions such as temperature and humidity.
- **Smart traffic:** It inculcates augmenting the transportation flow using transportation indicators, like the number of vehicles and pedestrians apart from informing travelers informing them about the next bus or train arrival or departure time using digital displays.
- **Smart industry:** It involves the enabling tracking of transport and logistics flows in a much easier way and not only for one industry but for different types of industries.
- **Smart energy:** It optimizes the urban energy systems and improving the quality of life for the citizens.
- **Smart devices:** A smart device is an electronic gadget that is able to interact with its user and is connected to other smart devices via networks. The most common smart devices are smart phones, tablets, smart watches and smart glasses. Smart

devices are rapidly emerging as popular appliances with increasingly powerful computation and processing. The difference between a smart device and those traditional ‘non-smart’ is the ability of smart devices to incorporate third party applications from different online markets. Smart devices can act as a person and can represent persona just like any other man in a conversation i.e., recognize speech and voice, react to commands, answer the questions and many more [9].

## 5 Futuristic Smart Devices

- **Nest Thermostat/Honeywell WIFI Smart Thermostat:** Both these thermostats enable you to control the temperature of your house from anywhere in the world using your smart phone. Additionally, both of these thermostats will adjust the temperature according to the user’s behavioral preferences [10].
- **Rain Machine:** The Rain Machine by Green Electronics attaches to your sprinkler controller and uses Wi-Fi to download seven days of weather forecasting, which means it can also work even when the WIFI is not working, but only for a few days. The sprinklers will remain off while raining.
- **Wally Home:** This incredible smart device can detect water leaks, and mold instead of vegetation. It comes with sensors that you can place anywhere in your house to monitor temperature, moisture and humidity. The sensors send you an alert when they detect leaks.
- **Vessyl Cup:** It is not just a cup. A smart revolutionary device analyses the liquids inside your cup and tells you what you are drinking including the amount of sugar, protein, calories fats, caffeine and other nutritional information.

## 6 Contribution of Smart Devices for Smart City Development

A smart city is generally a city equipped with smart technologies. However, smart technologies are implemented on smart devices and products with communicate to create a network which eventually on a larger scale make a smart city. The importance of smart devices is in the development of smart city development. Considering the smart living on a smaller scale i.e., a home or house (any location) equipped with smart technologies and devices like smart lighting system, smart security systems, entertainments and other things will make a smart home [11]. Now considering it on a little larger scale i.e., suppose buildings or blocks in an area equipped with the smart technology will make that area smart and the lives of the people living in smart living. Therefore, it will not be wrong to say that smart devices make technologies, homes and buildings, roads, and blocks smart. In addition, combining all the houses, buildings, blocks, roads, and areas will make a smart city. Some of the components

for developments of a smart city and the role of smart devices for fulfilling those components is being discussed below:

- **Smart Lightning Devices for Energy Conservation:** Smart lightning devices helps in day light harvesting and saving energy in case of no occupancy. This will help in unnecessary wastage of energy and later use in other important projects or even on overall production [12].
- **Safety and Security:** Smart devices like security cameras, alarming devices, etc. makes city more secure. For example, surveillance cameras that hold recordings of the streets and stores helps in catching some criminals and alarming devices can save unauthorized or forceful entries in one's house or even in government building or even office buildings [13].
- **Ease of Living:** Smart devices that controls temperature and ventilation of a house or building makes that place comfortable for living. The temperature, the air quality, the lightning can be adjusted according to one's comfort [14] .
- **Entertainment:** Smart devices like amazon Alexa, echo, or even robots are entertaining the end user or user can listen to songs, audio books or even news on just one voice command from any room [15] .
- **Information and Knowledge:** There are smart devices that can get any information over voice command. From listening news to finding the meaning of any word, everything just over a command [16] .
- **Healthcare and Emergencies:** Smart devices are implemented in emergency vehicles and hospitals and other places as police stations, fire stations etc. to provide 24/7 service [17].

## 7 Conclusion

This article has presented some latest trends and advancements in smart city development. The ultimate goal of a smart home system is to make the lives of the elderly and disabled more energy efficient. At the present time the services provided by smart home systems are reasonably priced. Consequently, the acceptance rate of utilizing a smart home system is increasing gradually. Smart home offers more security during your absence. Conceivably, the most intriguing future of research is the movement of this study to be used in more complex ecosystem of a big city and its inhabitants. The main goal of this study was to produce a reliable method and predictable resource without being able to measure a large city to produce practical information that would assist in the management and advancement of smart new cities. The purpose of this article was to provide information about the technological advancements and the implementation of that technology in our daily life. It has discussed numerous smart city constituents and how they have made living easier in a smart metropolis.

## References

1. I. Lopatovska, K. Rink, I. Knight, K. Raines, K. Cosenza, H. Williams, A. Martinez, Talk to me: Exploring user interactions with the Amazon Alexa. *J. Librarianship Information Sci.* (2018)
2. V. Fabi, G. Spigliantini, S.P. Cognati, Insights on smart home concept and occupants' interaction with building controls. *Energy Proced.* **111**, 759–769 (2017)
3. F. Mosannenzadeh, A. Bisello, R. Vaccaro, V. D'Alonzo, G.W. Hunter, D. Vettorato, Smart energy city development: A story told by urban planners. *Cities* **64**, 54–65 (2017)
4. F.K. Aldrich, Smart homes: past, present and future. In *Inside the Smart Home* (Springer, London, 2003), pp. 17–39
5. M. Lacinák, J. Ristvej, Smart city, safety and security. *Procedia Eng.* **192**, 522–527 (2017)
6. C. Nicolas, J. Kim, S. Chi, Quantifying the dynamic effects of smart city development enablers using structural equation modeling. *Sustainable Cities Soc.* **53**, 101916 (2020)
7. D.O. Aghimien, C. Aigbavboa, D.J. Edwards, A.M. Mahamadu, P. Olomolaiye, H. Nash, M. Onyia, A fuzzy synthetic evaluation of the challenges of smart city development in developing countries. *Smart and Sustainable Built Environment* (2020)
8. C. Manville, G. Cochrane, J. Cave, J. Millard, J.K. Pederson, R.K. Thaarup, et al., Mapping smart cities in the EU (2014)
9. B. Singh, Smart city-smart life-Dubai Expo 2020. *Middle East J. Business* **10**(4), 49–52 (2015)
10. A. Nikitas, K. Michalakopoulou, E.T. Njoya, D. Karampatzakis, Artificial intelligence, transport and the smart city: Definitions and dimensions of a new mobility era. *Sustainability* **12**(7), 2789 (2020)
11. H.H. Khan, M.N. Malik, R. Zafar, F.A. Goni, A.G. Chofreh, J.J. Klemeš, Y. Alotaibi, Challenges for sustainable smart city development: A conceptual framework. *Sustain. Dev.* **28**(5), 1507–1518 (2020)
12. A. Datta, New urban utopias of postcolonial India: 'Entrepreneurial urbanization' in Dholera smart city. *Gujarat. Dialogues in Human Geog.* **5**(1), 3–22 (2015)
13. Z. Wei, J.A. Landay, Evaluating Speech-Based Smart Devices Using New Usability Heuristics. *IEEE Pervasive Comput.* **17**(2), 84–96 (2018)
14. J.N. Arensmeier, P. Layton, U.S. Patent No. 9,741,023. (Washington, DC, U.S. Patent and Trademark Office, 2017)
15. C. Patrão, P. Moura, A.T.D. Almeida, Review of Smart City Assessment Tools. *Smart Cities* **3**(4), 1117–1132 (2020)
16. A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, Internet of things for smart cities. *IEEE Internet Things J.* **1**(1), 22–32 (2014)
17. D. Szpilko, Foresight as a tool for the planning and implementation of visions for smart city development. *Energies* **13**(7), 1782 (2020)

# Malpractice Detection in Online Assessments Using Eye Gaze Tracking and Object Detection



B. Nagaraju Yadav and M. Kameshwara Rao

**Abstract** Internet-based client authentication protocols must be strengthened to reduce attacks and security vulnerabilities that threaten the performance of apps in fast internet distribution and cloud computing. Due to a multitude of benefits such as effectiveness, convenience, simplicity, and usability, distance and digital training (called e-learning) seems to have become the mainstream in skills and retraining. Secondly, because just like the COVID-19 pandemic's physical isolation rules, online learning has now become the exclusive possibility. Due to the lack of physical existence, however, online systems are a major issue in monitoring attendees and students over sessions, particularly during tests. It is necessary to establish technological tools that deliver survey clearly for monitoring unfair, unethical, and unauthorized behavior in classes and examinations. In this dissertation, we develop a modern online proctoring system based on machine learning.

**Keywords** Online proctoring · Face detection · Eye gaze tracking · Object detection

## 1 Introduction

Mostly colleges as well as universities suggest additional coaching & courses in physical classrooms and regions, with lectures, enrollment exams, semester exams, and other activities requiring attendance. Intractability for students, teaching staff, as well as other professionals, the requirement of physical environments to strict standards, ease of access complexities for students and teachers in time and space, obstacles associated with individual learning difficulties, higher financial implications, travel obstacles and adversely affects to individuals and the environment, and

---

B. N. Yadav (✉) · M. K. Rao

Department of Electronics and Computer Engineering, KLEF, Vaddeswaram, India  
e-mail: [nagarajubayanaboina@gmail.com](mailto:nagarajubayanaboina@gmail.com)

M. K. Rao

e-mail: [kamesh.manchiraju@kluniversity.in](mailto:kamesh.manchiraju@kluniversity.in)

many others are among the disadvantages of classroom instruction in physical spaces. Most academic institutions offer educational courses and training in physical classrooms and regions, with lectures, written test, semester exams, and other activities requiring attendance [1]. The COVID-19 contagion has wreaked havoc on a variety of aspects of our way of life. Governments across the globe have controlled physical interaction for education, employment, and recreation in order to reduce human infection rates [2]. Education, as well as many other physical activities and organizations, have already been driven to relocate from manual to computerized places as a result of these events [3]. The rest of this paper is organized as follows. The research on online proctoring systems is explained in Sect. 2. The proposed system's methodology and design are explained in Sect. 3. The system is tested in Sect. 4. Results and Analysis.

## 2 Literature Review

Artificial intelligence (AI) has transformed our world and environs by bringing intelligence to almost all of our daily tasks [4, 5], but with a number of limitations [6, 7]. Machine learning (ml) have expedited innovations across a wide range of industries [8], including education, medical, transportation, communications infrastructure, disaster preparedness and intelligent transportation [9], to name a few. Intelligence refers to the ability to revolutionize web-based learning and invigilating in every manner. This paper proposes an innovative online proctoring system that use deep learning to consistently proctor geographical boundaries without the requirements for a physical proctor's presence. Face recognition using the OpenCV face recognition algorithm has been one of the biometric approaches used by the system. Throughout the exam, the online exam encounters numerous challenges [10] discussed the online exam's various difficulties and suggested a solution that also included grouping client hostnames or IPs for a specific area and time, as well as biometrics technologies such as face recognition and fingerprints [9]. In, a characteristics accreditation architecture for such official questionnaire is described, which is based on a variety of challenging problems, such as main remarks, intimate details, and intellectual questions. introduced a multi biometrics persistent authentication protocol which comprised identification of face, voice, touch, mouse, and keystrokes. For researchers and developers, the online proctoring system has recently become a challenge.

## 3 Proposed Framework

**Web Enrollment:** To register students' faces, we utilized HTTPS protocol to access the student's webcam during registration and captured the students' faces, which we then stored in the database. For web design, we used the flask micro-framework.

**Web Proctoring:** Web proctoring could be difficult to do during online exam sessions. The foregoing is some of the challenges:

- An unregistered student may take a test.
- The test could be taken by many students in the class.
- For face detection, the student may choose his still picture.
- During the test, the candidate could also use a smartphones, laptop, or ipads to stream a videos for facial recognition as well as books on a gadget like a smartphone, laptop, or iPad. In our principal objective is to alleviate those obstacles. We use biometric techniques like human activity recognition with gaze detection in the internet-based computer-based training control unit. The recommended system's algorithm is outlined in Technique is proposed. As a result, we use object recognition technologies like MobileNet SSD, which further help to avoid exam fraud by allowing students to use gadgets. The same MobileNet SSD model was used to identify the books and also the fare and secured test.

#### A. COCO Dataset

COCO Dataset is a dataset created by COCO. COCO stands for Common Objects in Context and refers to the picture collection that was established with the goal of improving image recognition. With the preponderance of the datasets containing modern neural networks, the COCO dataset includes demanding, next visible datasets for laptop vision. COCO is commonly used to determine the overall performance of real-time object detection systems, for example. The format of the COCO dataset is mechanically processed by advanced neural community libraries. The MS COCO dataset, which is a large key identification, segmentation, and captioning dataset, was released by Microsoft. For several laptop imaginative and prescient applications, the COCO dataset is widely used by gadget studying and laptop imaginative and prescient specialists. Computer vision aims to comprehend visual environments by recognising what objects are present, locating them in 2D and 3D, determining their properties, and defining relationships among them.

#### B. Mobilenet SSD

The cellular net-ssd version is a Single-Shot multibox Detection (SSD) community that identifies objects. It's an item detection version that determines an item's bounding enclosure and class from an input image. This version of the Single Shot Detector (SSD) item data at a particular Mobilenet as the spine and can achieve quick item detection optimised for mobile devices. The MobileNet technique employs a deep convolution operation as the basic layer network, which reduces the latency while keeping the model size in mind, but the accuracy for target recognition remains weak. Object recognition system In most instances, it is speedier than Faster RCNN. In this post, I'll go through what item detection is, what the TensorFlow API is, what the notion behind neural networks is, and how SSD structure works. Then I'll demonstrate you how to use TensorFlow API to enforce SSD MobilenetV2 educated over COCO dataset. You may find any one elegance from the instructions delivered

by the COCO dataset in this academic. After that, I assume you could implement your own SSD with maybe a little patience.

### C. Eye Gaze Tracking

We'll require a facial expression essential points decoder that can determine pupils in instantaneously for the purpose of activity recognition. We'll utilise the dlib module's for this which existing before network, which can discern '68 vital locations.' Dlib is implemented because, unlike a CNN Architecture, it can compare to the predicted in real time. The arithmetic arrangement of pixels around in the face is depicted in Fig. 1. Before we can proceed to image enhancement, we must first detect pupils. In order to locate the eyes, we must first recognize a face. A rectangular objects from of the dlib modules, which really is essentially a collection of facial dimensions, is required as input for the facial vital point detector.

### D. Object Detection

To categorise humans and smart phones with in webcam stream, we used the well prior weights of MobileNet SSD built just on COCO dataset. Using a pre-trained MobileNet SSD, we counted people on a camera. A warning can be triggered if somehow the count is not equal to one. The COCO dataset's mobile phone index is 67, thus we need to see whether any class indices are equal before reporting a smartphone. Because dlib is used for both eyetracking and mouth detection, we may create a separate threads for each, as well as a separate thread for the MobileNet SSD operations of people relying and motion detection. In the count \_people and phone() function, MobileNet SSD is applied to the camera stream. The categorisation of the things observed are then checked, and if many individuals or a mobile phone are found, corrective action is taken.

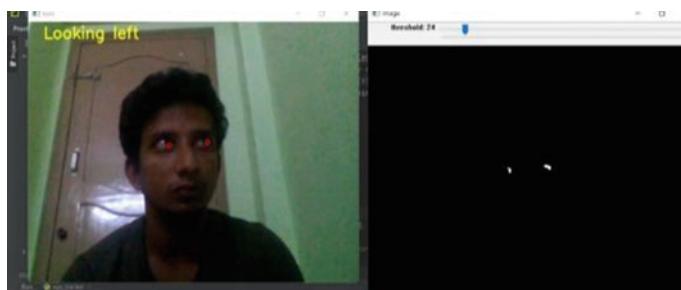
**Fig. 1** Eye gaze tracking  
(facial landmarks)



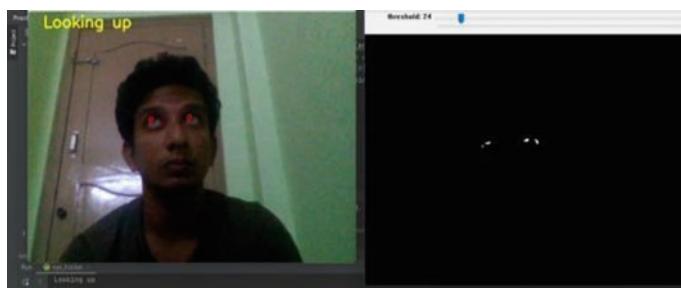
## 4 Results and Discussion

### I. Eye Gaze Tracking

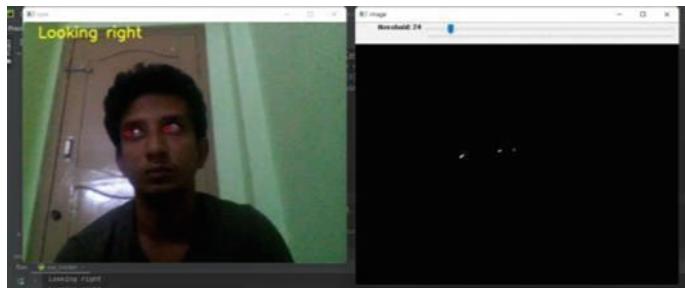
Real-time face Continuous monitoring, for example, is one of the empirical models, Object Detection undertake well with the proposed automated proctoring system. The threshold parameter for Eye Gaze Tracking must be adjusted in accordance with the illumination in the space People pulled the methodology to either the assignment employing numerous threshold values and discovered that when the threshold is set between 24 and 130, the model clearly recognizes the eyeball. display the model's results for recognizing eye gazing to the left and up, respectively (Figs. 2, 3, and 4).



**Fig. 2** Eye gaze tracking (looking left)



**Fig. 3** Eye gaze tracking (looking up)

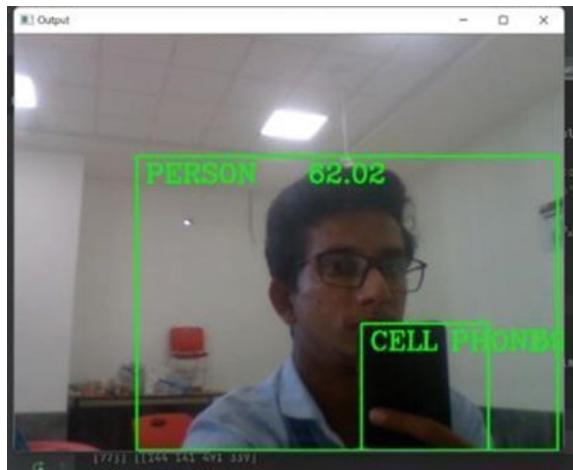


**Fig. 4** Eye gaze tracking (looking right)

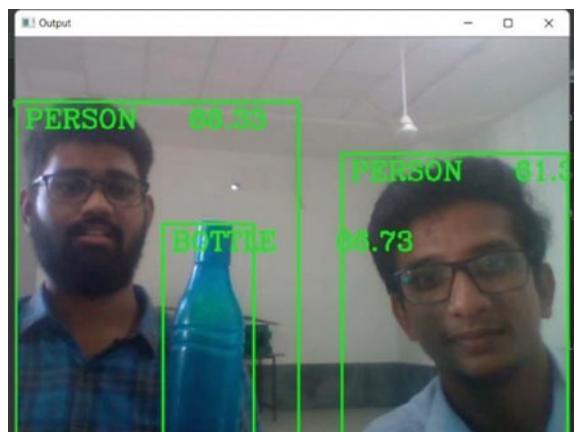
## II. Object Detection

We will learn how to detect objects using OpenCV and Python in this project. We'll employ the Object Detection OpenCV approach, that achieves a perfect balance between speed and accuracy. We will be able to execute this in real time with a high degree of accuracy, unlike many existing object detection approaches. To detect the face, we first must import the sample. We need the coco names dataset to train the data to detect objects, as well as weights and configuration files. We're going to import the files from of the mobile SSD we're employing. It will be expressed in nonsense. We have a single file with the architecture or settings. These are essentially available on the open CV documentation as of the beginning of 2020. The results are shown below in the Figs. 5, 6, and 7.

**Fig. 5** Object detection  
(mobile phone)



**Fig. 6** Object detection (bottle)



**Fig. 7** Object detection (mobile phone, person)



## 5 Conclusion

This study employs face recognition and object recognition techniques to provide complete knowledge for online assessments. During the online test, our proposed strategy will help decrease inequity. When using an online proctoring system, human-induced detection is crucial since it aids in spotting students' strange activities during the exam. In with us proposed approach, it doesn't include segmentation of economic activities scheme, consequently concentrating along with basis of certain characteristics authentication method technology and computer vision applications strategies for more than just the programme of available on the internet great deeds. In the future, plan to explore and critically evaluate a wide range of physiological behaviours including some looking out whatever screen door, speaking with people, going through different sorts of places, looking to move somewhere between, and

everything else. Even though there are various different object identification systems available, this was primarily employ the Mobilenet SSD model due of its quicker object detecting algorithms. In the research study, will examine at such approaches and compare them to the approach that had proposed so far.

## References

1. A. Wahid, et al., in *Toward Constructing a Secure Online Examination System*, Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication (2015), pp. 1–8
2. E. Alomari et al., COVID-19: detecting government pandemic measures and public concerns from Twitter Arabic data using distributed machine learning. *Int. J. Environ. Res. Public Health* **18**(1), p. 282 (2021) [Online]. Available: <https://www.mdpi.com/1660-4601/18/1/28>
3. B. Poutre et al., Combining testing software, online proctoring and lockdown browsers to assure a secure assessment environment for students in hybrid or online programs (poster 13) (Creighton University, Office of Academic Excellence and Assessment, 2015)
4. T. Yigitcanlar et al., “Can building” artificially intelligent cities safeguard humanity from natural disasters, pandemics, and other catastrophes? An urban scholar’s perspective. *Sensors* **20**(10), 2988 (2020) [Online]. Available: <https://www.mdpi.com/1424-8220/20/10/2988>
5. A. Chianese et al., An associative engines based approach supporting collaborative analytics in the Internet of cultural things. *Future Generat. Comput. Syst.* **66**, 187–198 (2017)
6. F. Alam et al., Analysis of eight data mining algorithms for smarter Internet of Things (IoT). *Proc. Int. Workshops Data Mining IoT Syst. (DaMIS)* **98**, pp. 437–442 (2016)
7. T. Yigitcanlar et al., Green artificial intelligence: towards an efficient, sustainable and equitable technology for smart cities and futures. *Sustainability* **13**(16), 8952 (2021) [Online]. Available: <https://www.mdpi.com/2071-1050/13/16/8952>
8. I.Y. Jung, et al., Enhanced security for online exams using group cryptography. *IEEE Trans. Educ.* **52**(3), 340–349 (2009)
9. E. Winarno, et al., *Anti-cheating Presence System Based on 3WPCA-Dual Vision Face Recognition*, 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (ECSI), pp. 1–5 (2017, September)
10. G. Apampa, et al., User security issues in summative E-assessment security. *Int. J. Digital Soc.* (2010)

# Big Data f5 Load Balancer with ChatBots Framework



**Kapil Joshi, Mohit Kumar, Minakshi Memoria, Pooja Bhardwaj, Gunjan Chhabra, and Dev Baloni**

**Abstract** Phase computing is the gathering of computer resources from several locations in order to accomplish a shared goal. The grid is a distributed system that handles interactive workloads with a high number of files. Grid computing offers a wide variety of geographically distributed computing resources and as a single resource that can be shared by any computer application for solving large-scale solutions, some planning algorithm that can control the different needs of the grid environment is required. Therefore, a planning algorithm for the phase environment is required, considering the load balance for fault tolerance in the phase environment. A chatbot is a humanlike conversational character, machine conversation system that interacts with user using natural languages. This system analyzes conversation through auditory or textual methods. We highlight pertinent topics and how these difficulties have been addressed utilising AI techniques after providing some context. The main purpose of chatbot is to control the entire communication structure to and from the boot (web service), the bots allows all type of client to communicate to the server. Also stimulates informal chat between a human user and a computer. We expect that through the years that chatbot will develop into real virtual human. People basically require an easy and efficient language to interact with the computers in the same way as people use common language for human communication. In this paper, proposed method gives better output with chat oriented dialog system and its technology to verify a proposed framework required for web service (bots).

---

K. Joshi (✉) · M. Memoria · D. Baloni  
Department of CSE, UIT, Uttarakhand University, Dehradun, India  
e-mail: [kapilengg0509@gmail.com](mailto:kapilengg0509@gmail.com)

M. Kumar  
Department of CSE, Chandigarh University, Mohali, India

P. Bhardwaj  
Department of CSE, Ajay Kumar Garg Engineering College, Ghaziabad, India  
e-mail: [bhardwajpooja@akgec.ac.in](mailto:bhardwajpooja@akgec.ac.in)

G. Chhabra  
Department of Systemics, School of Computer Science, UPES, Dehradun, India  
e-mail: [gchhabra@ddn.upes.ac.in](mailto:gchhabra@ddn.upes.ac.in)

**Keywords** BIG-IP f5 · Calculation phase · Data phase · Load balance

## 1 Introduction

Before there were computers, we could distinguish person from a non-person on the basics of their ability to participate in conversation. “But now, We have a HYBRID operating between person and non-person with whom we can talk in ordinary language. This system is developed as a [1] hybrid architecture that combines many methodologies in order to mimic human cognitive features. The building of an intelligent machine with conversational ability is an ambiguous goal here. However, because it includes language comprehension and discourse management, putting this into practice is a difficult process.

A conversational agent is a chatbot that interacts with a user based on a PATTERN MATCHING Mechanism between user queries and a set of well-defined knowledge base rules. Different human dialogue system, i.e., Chatbot are developed using set of conditions like topic, goal, speech etc. Starting from ELIZA (Weizenbaum 1960) that stimulates Psychotherapist, and then PARRY (Colby 1973) which stimulates paranoid patient. In recent years, AI research has seen the rise of the Artificial General Intelligence (AGI) paradigm, which focuses on the learning process before formalizing domains. We tweaked Wallace’s ALICE (Artificial Linguistics Internet Computer Entity) implementation from 1995.

Grid computing provides an environment for sharing and accessing numerous computing and storage resources that are diverse and geographically distributed. There are two types of phases based on its operation—the computational phase and the data phase. In the computational grid, a large [2] computational task is carried out by different machines in parallel and returned to the original machine as a result. In Data Grid, different machines or different people on the same network can access data, share data to perform updates and some other functions.

Load balancing is a technique used to spread loads across multiple nodes so that maximum efficiency is used. It is a networking method for distributing loads across multiple computers or multiple resources in a single cluster or multiple clusters. Maximize resource utilization, maximize performance [3] and usability and maximize response time to successful load balancing. Some existing load balancing methods are rapid adaptation, bee-inspired, heat dissipation-based, and dynamically structured P2P systems [4].

## 2 Background

In the 1950s and 60s, the famous computer scientist **ALAN TURING** gave the concept in terms of computer, i.e., Concept of TURING MACHINE, to test the computer ability to display intelligent behavior equivalent to that of a human. ELIZA,

**Fig. 1** An introduction to chatbots



the first chatterbot in computer science history was invented by JOSEPH WEIZENBAUM in 1996; and seemed to be fool users into believing that they were conversing with a real human. PARRY (1972) is the another classic chatbot developed by the psychiatrist Kenneth Colby. PARRY has been initially designed to perform as a training tool for psychiatry students. In 2016, the chatbot were considered as one of the hottest trend in technology. Chatbot systems were supported by biggest messaging platforms like Facebook, Messenger, Skype in Fig. 1.

### 3 Architecture for Chatbot

The architecture is built on the AGI paradigm: a modular, easily controllable, and upgradable architecture that blends symbolic and sub-symbolic reasoning capabilities. Chatbot's architecture combines a language model and computer techniques to simulate natural language informal interaction between humans and users. The system architecture is made up of various components.

#### 3.1 *Dialogue Engine*

The Dialogue Engine is made up of a number of modules that determine a chatbot's knowledge and reasoning capabilities as well as how the user and chatbot communicate. It improves the standard ALICE Dialogue mechanism. It establishes the principles for lexical [5] management and sentence comprehension. The following are the primary disadvantages of this strategy: (a) It is time consuming. (b) A dialogue mechanism that is too rigid.

### 3.2 Chatbot Module

Each module in the chatbot framework has its own specific feature that makes it unique. For example: We do differentiation on functions, topics, emulation on specific goals etc.

## 4 The Convergence of Bots Platform

Here, we discuss about some of the messaging platform where the bots can be built up. Facebook messenger are the most popular messaging platform with more than 1 billion monthly active users.

### 4.1 The Optimizer

This is the largest category of *Functional Bot*. These bots are those who take concrete challenges and try solving it; better than the existing applications or websites. For example: If we have some electric appliances like A.C, in the living room, then we can simply switch it on by saying, “Switch on A.C.”. Instead of clicking number of buttons. But the huge disappointment was felt when these bots failed to fulfill their promises.

### 4.2 The One-Trick Pony

This refers to that if we want to explain a person with his face that he has a naive smile, smoky eyes, long beard [6] etc. Then we can simply say “A Frank Assessment of his face.”

It relates that if a thing has many attributes or want to display a lot of messages then with the help of One-Trick Pony, We can easily display it with images or videos.

### 4.3 Proactive Bots

They are the one who message the customer first with below categories.

**Fig. 2** User interaction with the help of natural languages



#### 4.3.1 The Social Bots

Super interesting bots which are meant to accomplish a task and the most important feature of this bots is that they develop on the power of group or unity. They usually make some unique nature of messaging [7] platforms.

#### 4.3.2 The Shield Bots

These are similar to optimizers in their function. The main idea of this bots is to exchange a talk of a cold hearted person with a friendly robot. These bots includes customer support issues, booking a railway ticket, parking etc. and exist till now due to its ability to do work according to their alternatives.

#### 4.3.3 The Super Bots

One of the intelligent bots which is a chatty and social bots and it embraces the advantages of the messaging canvas. It provides natural language responses to your spoken input, exchange information with other apps and web services, control the activation of device features, generates personalized responses for the ordinary conversation. Examples- Alexa, Siri, Cortana and Allo in Fig. 2.

### 5 Human to Human Versus Human to Chatbots Dialogue

According to research, GOOGLE made a chatbot that debates the Meaning of Life. Here, a human talk with machine and conversation starts in this manner.

The machine learned the way of conversing the enormous collection of old movie dialogues. Google also detailed that chatbot can do reasonable conversation like

**Fig. 3** Human to human versus Human to Chatbot dialogue



humans and with humans and also concluded that they let the machine learn from data rather than the rules of hand-coding.

According to Perlis 1990, chatbots demonstrate a poor capacity to reason about conversation, and also cannot identify misunderstanding and generally talk at superior level. But, According to Reinie 2009, chatbots provide the most relevant output according to the intermediate input in Fig. 3.

## 6 Chatbots as a Tool to Visualize or Animate Corpus Text

Chatbot tools are those computer program which allows a user to interact using natural languages. It is also considered that chatbot serve as a tool to amuse user through chatting with a robot. Here, a program is presented to convert a machine readable text (Corpus) to a specific format (chatbot) which is closer to human language. Different Corpora were used: Dialogue Corpora, which is a monologue corpora where questions and answers are in pair of turn. The general goal is to use text-like sources to bootstrap language knowledge and resources. Leeds University has conducted a wide spectrum of research during the past 20 years, which has been grouped into various sub-topics like constituent-lielihood statistical and modeling of English Grammar mentioned in Fig. 4.

Using AIML knowledge, development of a Java program was done that converts a readable text (corpus) to the AIML format. After reading the text, it was inserted in a vector. Then, the vector was passed through many reprocessing phases where all the linguistics annotation was filtered. After that the entire module was divided into two parts: First is pattern and second is template.



**Fig. 4** Messaging based platform

## 7 Method of Load Balancing Using f5 with Chatbot

We have proposed a method for load balancing using BIG-IP f5. It uses round robin scheduling to plan jobs at different ends. If something goes wrong at one end, the work will run on the other end using this technique. So this will give us fault tolerance. We have presented a load balancer system using BIG-IP f5 that implements a round robin scheduling to plan jobs at different ends with different chatbots or human computer dialogue system. Computer dialog system were developed starting with ELIZA then PARRY. We presented a tool that explores and visualize different types of english language used in BNC corpus in a contrast way and provides a qualitative illustration of the language of a specific speaker type or domain.

Future research includes the main focus on learning process then on formalization of domain and also about the ALICE style chatbots that can be used as English conversation practice tool. People probable wonder that “Why does anyone care about chatbot.” They look so simple to deal with and not a big task of dealing it with. But, the reason is that it’s because people in today’s time are using messenger application more than any other social application.

Here we provide a live example of a web server farm. The two different web servers we provide are managed by f5 (BIG-IP). For example Google, Yahoo etc. can handle millions of web requests per f5 milliseconds. We have two web servers. There are: Op-Web and ABC-Web.

These web servers are managed by a web server pool. Who do not manage the server and monitor their health. If no server services are running, this pool will downgrade them and those servers will go into maintenance or restoration process. In this case the other servers will take place in the space of that server.

## 8 Result and Discussion

We have total five servers, In a One pool we have three servers attached & in another pool we have two servers. Server IP address are given below along with respective pools.

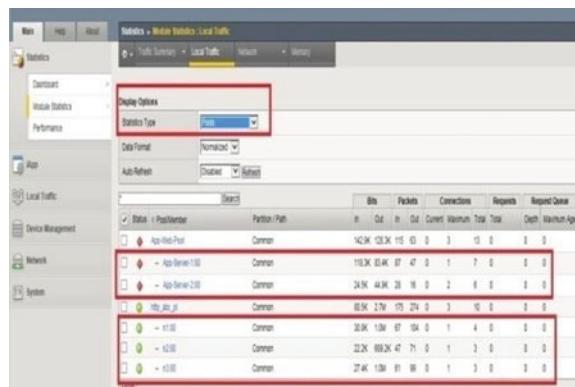
APP-WEB has following servers. These two servers are windows servers.

1. 10.2.0.8 Server1
2. 10.2.0.9 Server2

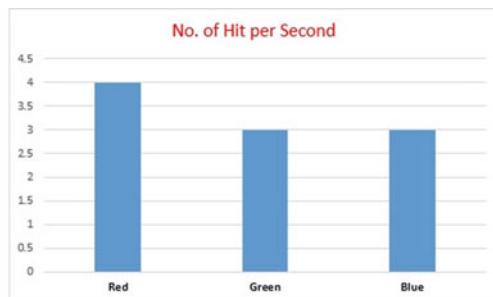
ABC-WEB has following servers. All three are Linux based servers in Figs. 5, 6, and 7.

1. 10.2.0.11 Red11
2. 10.2.0.22 Green22
3. 10.2.0.33 Blue33

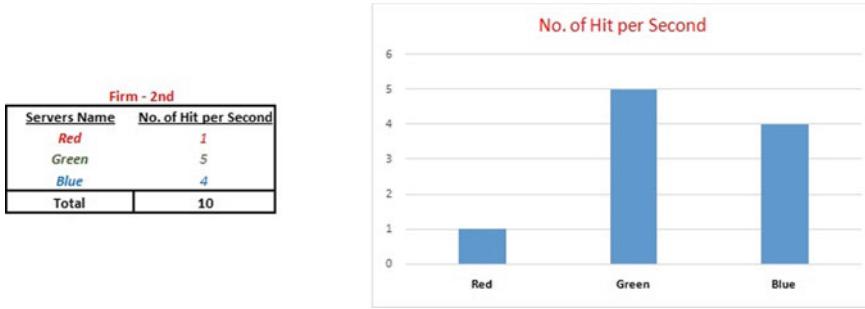
**Fig. 5** Shows the no. of request handled by each pool of servers and pool members of servers



Firm - 2nd	
Servers Name	No. of Hit per Second
Red	4
Green	3
Blue	3



**Fig. 6** Shows the no. of hits by second farm of servers



**Fig. 7** Shows no. of hits after occurring fault in red server

## 9 Conclusions

It concludes that this method gives maximum resource utilization even after occurring fault in system than existing algorithm. It gives maximum throughput and efficiency as this method is using round robin scheduling. According to the Fig. 6 and Fig. 7 we can compare that total hits that have been distributed among server green and blue after occurring fault in first red server. We have presented a fault tolerant system that implements work even occurrence of fault with a framework which enables user to interact with it using natural languages. Different chatbots or human computer dialogue system were developed starting with ELIZA then PARRY. In this paper, We present a tool that explores and visualize different types of english language used in BNC corpus in a contrast way and provides a qualitative illustration of the language of a specific speaker type or domain.

## References

1. N. Svenningsson, M. Faraon, in *Artificial intelligence in conversational agents: A study of factors related to perceived humanness in chatbots*. Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference (2019), pp. 151–161
2. M.H. Chao, A.J. Trappey, C.T. Wu, Emerging Technologies of Natural Language-Enabled Chatbots: A Review and Trend Forecast Using Intelligent Ontology Extraction and Patent Analytics. Complexity (2021)
3. K. Lyons, *Intelligent Chatbot: Technical Report* (Doctoral dissertation, Dublin, National College of Ireland, 2017)
4. L. Olsen, Can chatbot technologies answer work email needs?: A case study on work email needs in an accounting firm (2021)
5. K. Mabunda, *An Intelligent Chatbot for Guiding Visitors and Locating Venues* (2020)
6. S. Li, H. Zhang, Z. Jia, C. Zhong, C. Zhang, Z. Shan, et al., Understanding and addressing quality attributes of microservices architecture: a systematic literature review. Info. Soft. Tech. **131**, 106449 (2021)
7. P.H. Beaumont, *Digital Finance: Big Data, Start-ups, and the Future of Financial Services* (Routledge, 2019)

# Keylogger Awareness and Use in Cyber Forensics



Anuj Kumar, Kamlesh Kumar Dubey, Himanshu Gupta, Sumita lamba, Minakshi Memoria, and Kapil Joshi

**Abstract** Today hacking is not far enough from the reach of common people. But still today people are using untrusted systems (e.g. Cyber cafés, Internet labs in organizations etc.) to deal with their important financial, confidential and password accounts, where Keylogger is playing a vital role for these cyber criminals to peep into anyone's account. Keylogger, Keystroke tracker or KeyRecorder is software which records the key pressed and stores them in secondary memory storage. This hack-attack is widely used in today's era but still unfortunately some people are unaware of this technology. Not only Keylogger is a threat but Screen Recording software is also one of the severe attacks, to steal the password of the user which is a counter attack to the use of virtual keyboards. This paper focuses on public interest in such security training. This paper is divided into two phases. Phase I deals with dark side of Key Loggers i.e. Key Logger as a Hacking tool and Phase II deals with the silver side i.e. Key Logger as a Security tool. At last there will be some tricks that will help to keep our data away from such Keylogging softwares.

**Keywords** Anti-screenshot technology · Case studies · Online Banking · Screen recording software and virtual keyboard

---

A. Kumar · M. Memoria (✉) · K. Joshi  
Department of CSE, UIT, Uttarakhand University, Dehradun, India  
e-mail: [minakshimemoria@gmail.com](mailto:minakshimemoria@gmail.com)

K. K. Dubey  
Department of Applied Science and Humanities, Invertis University, Bareilly, India

H. Gupta  
Department of CSE, College of Engineering Roorkee, Roorkee, India

S. lamba  
Department of IT, Ajay Kumar Garg Engineering College, Ghaziabad, India  
e-mail: [lambasumita@akgec.ac.in](mailto:lambasumita@akgec.ac.in)

## 1 Introduction

Keylogger are the software or hardware devices which record each and every key pressed by the user and store them. These keyloggers are used almost in every web hack-attack, mostly to steal bank account passwords of the users. In the world of eBanking, e-Commerce and e-Mailing [1], it has become a slaughterers threat that how to prevent our passwords from cyber criminals [2]. Moreover the threat has taken a severe form after the introduction of screen recording software which paralyzes the anti keylogging [3]. Though Internet has made banking and other account managing things easy for us but on the other side it has also provided insecurity to our belongings. Not only keylogger are softwares but also there are hardware keyloggers which look like a USB and may get attached to your CPU and steal your passwords. Such hardware Keyloggers are even more dangerous as they are least or 0% detected by software means (Ant-Virus etc.'s). KeyLoggers can track 2 million keystrokes (approx). Not only this was enough but Screen Recording softwares are no less in killing cyber security. These kinds of softwares take a screenshot of every screen that appears on desktop when the mouse key is pressed, giving out the password to the hacker.

## 2 Keylogger

**Software Keylogger** Software Keyloggers can be installed in victim's system by two means:

**Through Internet** These kinds of keyloggers can be purposely installed in victim's computer by sending him an e-mail attached with a Trojan that works as a keylogger. These files can be sent as e-greeting cards etc. or they can look like any file which may attract you and you would definitely install it on your system. On gaining access to your system, they get automatically installed and they do not even get showed in task manager or even detected by anti-virus softwares. So easily they record the victim's keystroke and send it to the attacker's email address and keep on updating him automatically if you have an access to internet.

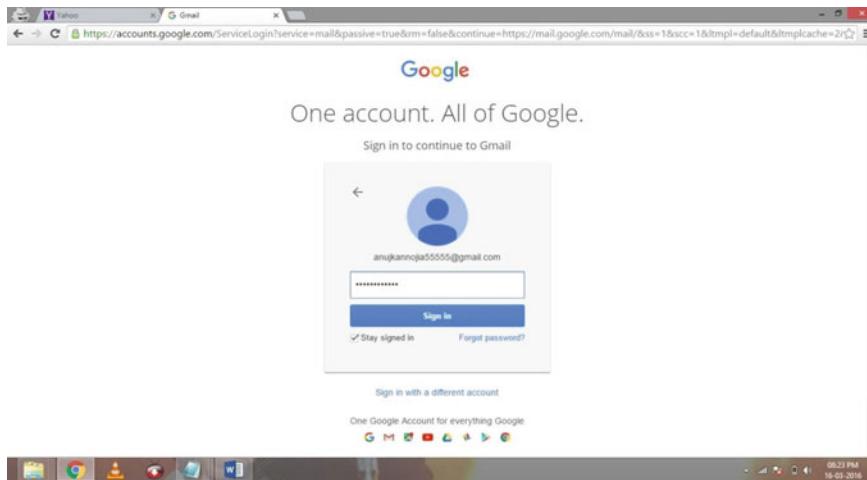
**Through physical access to your system** The attacker can get physical access to your untrusted system (Cyber cafes, public internet labs etc.) or even to your personal system and can install the software with a security password which victim can never detect. Many software keyloggers are easily available in the market. Figure 1 shows an example of a software keylogger:

Here I will show you an image of how a software keylogger can track your password. I will show you an example of a G-mail ID and its password (Fig. 2).

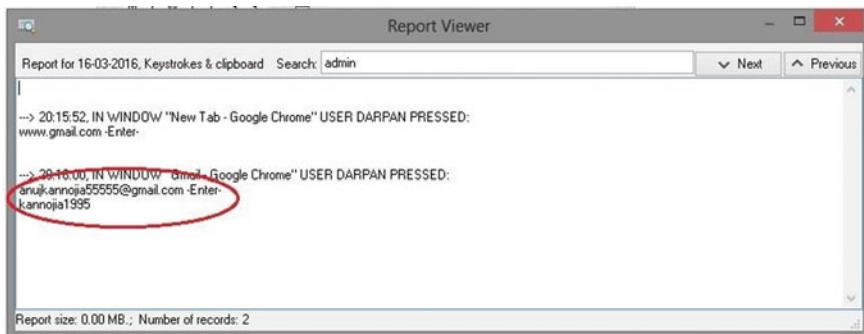
Another Fig. 3 shows the report generated by the keylogger, which shows the user's G-mail ID and password.



**Fig. 1** An example of software keylogger



**Fig. 2** An example of G-mail ID login page entering the ID and password. <https://free-keylogger.en.softonic.com/>



**Fig. 3** Report generated by software keylogger. <https://gmail.com/>

### 3 Proposed Study

Hardware keylogger are used for keystroke logging. They plug in between a computer keyboard and record all keystrokes, Chats, email passwords and other sensitive information. A hardware keylogger looks like a USB or other peripheral device so the victim can never doubt that it is a keylogger [4]. The hardware keylogger has an inbuilt memory which stores the sensitive information. Any attacker who can get physical access to your system can easily plug it in your CPU or fabricate it inside your keyboard to gain your secret information. Mostly hardware keyloggers are dangerous but due to their remoteless feature they are not widely used. They are generally used in *Untrusted* systems like Internet Café's, public internet booths etc. Figure 4 shows a Hardware keylogger:-[4]

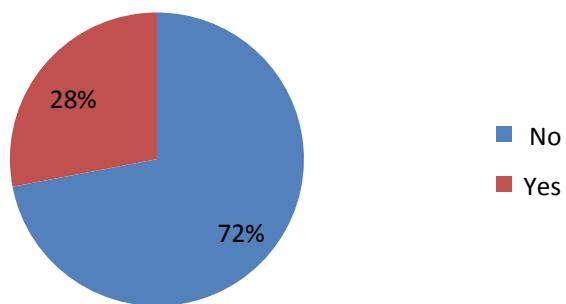
**Screen recording software** is such malicious programs that can take a screenshot on every mouse click. These screen recording softwares are a counter attack to Virtual Keyboard. Even if someone aware of Keylogger and uses Virtual Keyboard can easily get in trouble if he/she is not aware of such softwares [3]. What this Screen Recording software do is: that when the victim uses virtual keyboard, a keyboard appears on desktop screen whose keys are to be pressed by mouse. When the mouse key is pressed (that key may not be recorded by the Software/Hardware keylogger)



**Fig. 4** Hardware keyloggers

**Fig. 5** Unawareness of users about security training [6]

### Do users know about security training ?



it can be recorded by taking screenshot of that event just after mouse clicked which will take a snap of the key pressed in virtual keyboard thus giving out the password.

## 4 Public Awareness About Keyloggers

We have seen that users today are not at all aware about any security training. Not only stand alone users but roaming users are also not aware about any security measures. Though if they are aware, but least of them take interest in preventing their data from such hack-attacks [5]. From my research I got this result as shown in Fig. 5

Joe Lopez's losses were caused by a combination of overall carelessness and an ordinary keylogging program [6].

## 5 Security Measures to Prevent Keylogging

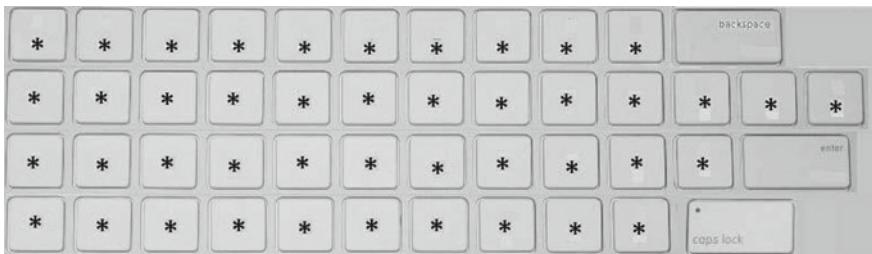
**Virtual Keyboard** Virtual Keyboard is the implementation of a software based keyboard which may help in preventing security from Keyloggers (both hardware and software keyloggers). Virtual Keyboard makes the use of event triggered by the mouse click and in that click on virtual keyboard's key the password is feeded. Figure 5 shows the virtual keyboard:

Virtual Keyboard can provide 75% security from keyloggers but still cyber criminals have a tool to break this security passage i.e. Screen Recording Software.

1. **Anti Screenshot Technology** Screen recording software is that software that bypasses the security provided by virtual keyboard. So in order to prevent screen recording "Anti-Screenshot Technology" is used. The anti-screenshot technology is basically advanced virtual keyboard. When user uses these keyboards he/she roams the cursor on the anti-screenshot software. While moving the cursor on this keyboard the keys in that specific row on which cursor is pointed



**Fig. 6** Example of anti-screenshot software's virtual keyboard [2]



**Fig. 7** When a mouse click event has occurred [2]

turn into some symbols like asterisk (\*) or hash (#). When any virtual key is pressed at that time all the keys are turned into (\*) or (#), so what the screenshot will be of, is only the symbols not the keys. And after that all the keys of that anti-screenshot keyboard are shuffled in such a way that the Trojan will take millions of attempt to guess the right password but since the number of retries will be just three so the security will be maintained. No doubt that “Anti-Screenshot Technology” [2] provides a far better security to such threats. Here is the example of Anti Screenshot Software in Figs. 6 and 7.

2. Check the system (Untrusted system) for any extra hardware device. It may be a keylogger which may look like a USB.
3. Keep your anti-virus software up to date because many Anti-viruses today are coming with anti keylogging protection.
4. Don't get attracted to any e-greeting or such files on the internet which are doubtful.
5. Manage your account and always use filter evasion to protect your accounts.
6. Avoid *untrusted* systems to deal with important password logging accounts.
7. Always maintain some backup for your data if it is really important.
8. Use Anti-Screenshot Technology to deal with most crucial account logins. Though its efficiency is not as good as of hardware keyboard but all above it provides 99% security [2].

## 6 Conclusion

Since the cyber theft is increasing day by day so it is an urge to get aware of keyloggers. Online Banking is the main point of concern because various cyber thefts from different parts of the world in online banking are just due to keyloggers. Though keyloggers are widely used in cyber crimes but taking its brighter side, Keylogger can also be used as a security measure to have a check on activities going on our computer. For e.g. Parents can keep a check on their children's activity, a boss can keep an eye on his employee's activities etc. Not talking about our personal data but the data at International Security level is to be kept away from such threats.

After all, if you think your data is important and worth protecting, the chances are that someone else will think it is worth stealing.

## References

1. S. Gong, *Design and Implementation of Anti-Screenshot Virtual Keyboard Applied in Online Banking “E-Business and E-Government (ICEE)*, in 2010 International Conference (7–9 May 2010), pp-1320–1322
2. A. Parekh, A. Pawar, P. Munot, P. Mantri, Secure authentication using anti-screenshot virtual keyboard. Int. J. Comp. Sci. (2011)
3. N. Adhikary, R. Shrivastava, A. Kumar, S. Kumar Verma, M. Bag, V. Singh, Battering keyloggers and screen recording software by fabricating passwords. Int. J. Comp. Net. Info. Security (2012)
4. G.N.U. Logkeys-a, L. Keylogger, The source code for the index construction and search is available at <https://github.com/kernc/logkeys>. Implemented on C and dual licensed under the terms of either GNU GPLv3 or later, or WTFPLv2 or later (2021)
5. A. Wajahat, A. Imran, J. Latif, A. Nazir, A. Bilal, *A Novel Approach of Unprivileged Keylogger Detection*. In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (2019)
6. S. Ortolani, C. Giuffrida, B. Crispo, Bait your hook: a novel detection technique for keyloggers. In International workshop on recent advances in intrusion detection (Springer, Berlin, Heidelberg, 2010), pp 198–217

# Author Index

## A

- Abhyankar, D., 655  
Adithya Harish, S. M., 567  
Aditya Rao, 539  
Ajay Lakhnotra, 187  
Akanksha Bali, 63  
Akash Saxena, 289  
Alihodzic, Adis, 447  
Amato, Alberto, 263  
Amit Agrawal, 351  
Amit Mahajan, 53, 97  
Amit Sharma, 187  
Amrendra Tripathi, 663  
Aniruddh Sharan Mangla, 465  
Anju Singh, 197  
Anshika Jain, 283  
Anubha Jain, 77, 87, 113  
Anuj Kumar, 559, 663, 719  
Archana Barthwal, 663  
Aryan Khari, 519  
Asha, S., 623  
Ashok S. Sangle, 427  
Astha Pareek, 17  
Atul Kumar Ramotra, 97  
Atul M. Gonsai, 273  
Avinash Sharma, 229

## B

- Balakrishnan, S., 43  
Balasubramanian, M., 43  
Bharathi, N., 215  
Bhoomika Batra, 477  
Bindhu, M., 623  
B. Nagaraju Yadav, 701

## C

- Chattar Singh, 439

## D

- Darshil P. Patel, 419  
Deepali Bhende, 329  
Devanand Padha, 187  
Dev Baloni, 709  
Dhara Rathore, 289  
Dhiren B. Patel, 33

## E

- Ekta Narwal, 123, 205

## G

- Gaurav Kumar Das, 289  
Gopal Sakarkar, 329  
Gopirajan, P. V., 319  
Gunjan Chhabra, 709

## H

- Hanumanthappa, M., 77, 159, 215  
Hasan, Mohammad Kamrul, 167  
Hema Priya, N., 567  
Heta S. Desai, 273  
Himanshu Barhaiya, 237  
Himanshu Gupta, 719  
Hrosik, Romana Capor, 447

## I

- Ingle, M., 655

**J**

- Jatinder Manhas, 25, 103, 187  
 Jaya Saxena, 113  
 Jayshri Patil, 139  
 Jikitsha Sheth, 139  
 Jitender Sehgal, 663

**K**

- Kalavathi, B. N., 159  
 Kamlesh Kumar Dubey, 719  
 Kapil Joshi, 663, 709, 719  
 Kavitha, S., 159  
 Kheni Rukshmani, S., 419  
 Khushboo Saxena, 549  
 Kinjal A. Patel, 419  
 Kinjal N. Patel, 419  
 Krishna Mohan, 519  
 Kuljeet Singh, 53

**L**

- Laavanya Mohan, 579  
 Lecce Di, Vincenzo, 263

**M**

- Madhu Gaur, 407  
 Madhulika Bhatia, 361, 373, 465  
 Manik Bhan, 229  
 Maya Ingle, 251, 283, 591  
 Mehak Giri, 229  
 Mehak Khanna, 539  
 Minakshi Memoria, 559, 709, 719  
 M. Kameswara Rao, 615, 701  
 Mohit Kumar, 663, 709  
 Monalin Pal, 683  
 Monika Patel, 33  
 Monika Rathore, 673  
 Monika Sharma, 455  
 Muniasamy, Anandhavalli, 485  
 Muskan Gupta, 519

**N**

- Navdeep Singh, 387  
 Naveen Raju, D., 319  
 Navneet Sharma, 307  
 Neha Tiwari, 307  
 Nilesh Jain, 237  
 Niram, 123  
 Nirupama Ansingkar, 395  
 Nirupama Singh, 11  
 Nishit Dhamija, 529

Noorul Julaiha, A. G., 631

**P**

- P. Anitha Rajakumari, 509  
 Panneer Arockiaraj, S., 221  
 Pinal J. Patel, 419  
 Piuri, Vincenzo, 263  
 Ponduri Vasanthi, 579  
 Pooja Bhardwaj, 709  
 Popolizio, Marina, 263  
 Prabhavathy, M., 1  
 Pradeep, R., 319  
 P. Radha Krishna, 113  
 Prapti D. Deshmukh, 395, 427  
 Pratham Vats, 373  
 Pratima Sen, 251  
 Praveen Kumar, 361, 373, 465  
 Praveen, M., 623  
 Preeti Manani, 373  
 Preeti Tiwari, 387, 407  
 Pritee Parwekar, 341, 509  
 Priyanka Dewatwal, 11  
 Priyanka Mathur, 607  
 Priyanka Mitra, 131  
 Priyanka Verma, 455  
 Priya Rani, 103  
 Priya Sinha, 539  
 Priyatharshini, R., 631  
 Puninder Kaur, 229

**R**

- Radhika Mohan Gupta, 607  
 Rajanish Jain, 351  
 Rajashree V. Bothale, 113  
 Rajesh Shrivastava, 167  
 Rajkumar, Y., 43  
 Rajneet Kaur Bijral, 25  
 Rajwinder Kaur, 229  
 Ramachandran, N., 77  
 Ramnarayan, 559  
 Ravikumar, K., 77  
 Ravi Subramanian, N., 567  
 Renu Devi, 205  
 Retinder Kaur, 175  
 Rishi, O. P., 643  
 Rita B. Patil, 395  
 Ritesh Joshi, 251  
 Ritu, 123  
 Ritul Bakshi, 299  
 Roheet Bhatnagar, 485  
 Roopa, H. R., 221

Rubini, P., 683  
Ruchi Jain, 197  
Ruchi Nanda, 607

Surendiran, B., 1, 567  
Suresh Kaswan, 175  
Suresh Kumar, K., 43, 319  
Switek, Slawomir, 495

## S

Saksham Goyal, 361  
Salini Suresh, 77, 87  
Santhi, S. G., 615  
Saranya Rubini, S., 1  
Saru, 373  
Saru Dhir, 519, 529, 539  
Sasikumar, S., 319  
Saureng Kumar, 147  
Saveeth, R., 1  
Shaik Mohissin Sultana, 615  
Shallu Kotwal, 103  
Sharma, S. C., 147  
Shaurya Gupta, 693  
Shaveta Bhatia, 465  
Shilpa Singh, 17  
Shivani Dubey, 351  
Shivani Sharma, 643, 673  
Shruti Ahuja, 341  
Shukla, Vinod Kumar, 693  
Simar Preet Singh, 167  
Simran Kaur Hora, 591  
Skandha Choudary, 341  
Sonali Vyas, 693  
Sourabh Singh Verma, 693  
Sumeet Gill, 123, 205  
Sumita lamba, 719  
Suneeta, V., 87  
Sunil Ghildiyal, 559  
Sunitha, 87

T  
Tasleem Arif, 103  
Tuba, Eva, 447  
Tuba, Ira, 447  
Tuba, Milan, 447

## U

Umamaheswari, B., 131  
Uma Maheswari, S., 1  
Urmila Patel, 419  
Urvashi, 361

V  
Vaibhav Bhatnagar, 477  
Vaibhav Gandhi, 597  
Vibhakar Mansotra, 53, 63, 97  
Vijay Gupta, 439  
Vijay Singh Rathore, 131, 299  
Vijeta Kumawat, 131  
Vikas Singhal, 351  
Vimal Kumar, K., 341  
Vinod Sharma, 25

## Y

Yogesh Kumar Gupta, 549  
Yogita D. Bhajgotar, 597