Lecture Notes in Electrical Engineering 990

Sarika Jain
Sven Groppe
Nandana Mihindukulasooriya *Editors*

# Proceedings of the International Health Informatics Conference

## IHIC 2022

Springer

# Lecture Notes in Electrical Engineering

## Volume 990

**Series Editors**

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy
Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico
Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany
Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China
Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore
Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany
Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China
Gianluigi Ferrari, Università di Parma, Parma, Italy
Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain
Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany
Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA
Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt
Torsten Kroeger, Stanford University, Stanford, CA, USA
Yong Li, Hunan University, Changsha, Hunan, China
Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA
Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain
Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore
Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany
Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA
Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany
Subhas Mukhopadhyay, School of Engineering and Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand
Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA
Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan
Luca Oneto, Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy
Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi Roma Tre, Roma, Italy
Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Gan Woon Seng, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany
Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal
Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China
Walter Zamboni, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy
Junjie James Zhang, Charlotte, NC, USA

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

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

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

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

**China**

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

**India, Japan, Rest of Asia**

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

**Southeast Asia, Australia, New Zealand**

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

**USA, Canada**

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

**All other Countries**

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

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

Sarika Jain · Sven Groppe ·
Nandana Mihindukulasooriya
Editors

# Proceedings of the International Health Informatics Conference

## IHIC 2022

Springer

*Editors*
Sarika Jain
National Institute of Technology
Kurukshetra
Kurukshetra, India

Sven Groppe
University of Lübeck
Lübeck, Germany

Nandana Mihindukulasooriya
MIT-IBM Watson AI Lab
Cambridge, MA, USA

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-19-9089-2 ISBN 978-981-19-9090-8 (eBook)
https://doi.org/10.1007/978-981-19-9090-8

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

# Organization

## Patrons

Chief Patron: Rajita Kulkarni, Honorable President of Sri Sri University

Patron: Ajay Kumar Singh, Honorable Vice Chancellor of Sri Sri University

## Advisory Committee

Laxmidhar Behera, IIT Kanpur, India

Milos Stojmenovic, Singidunum University, Danijelova 32, Belgrade, Serbia

Aleksandar Jevremovic, Singidunum University, Danijelova 32, Belgrade, Serbia

George Tsaramirsis, Higher Colleges of Technology, Abu Dhabi, UAE

## Honorary Chairs

Rabinarayan Satpathy, Sri Sri University, Cuttack, Odisha, India

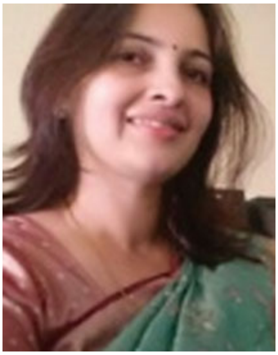
Sunil Kumar Dhal, Sri Sri University, Cuttack, Odisha, India

## General Chairs

Sarika Jain, National Institute of Technology, Kurukshetra, Haryana, India

Sven Groppe, University of Lübeck, Germany

## Track PC Chairs



Suneeta Satpathy, Sri Sri University, Cuttack, Odisha, India

Archana Patel, School of Computing and Information Technology (CIT), Eastern International University (EIU), Vietnam

Sudhir Kumar Mohapatra, Sri Sri University, Cuttack, Odisha, India

Bibhu Kalyan Mishra, Sri Sri University, Cuttack, Odisha, India

## Publication Chair

Jyotirmoy Chatterjee, Lord Buddha Education Foundation (Asia Pacific University of Technology & Innovation), Kathmandu, Nepal

Sachi Nandan Mohanty, Department of Computer Science, Singidunum University, Serbia

Rajesh Ojha, Sri Sri University, Cuttack, Odisha, India

## Publicity Chair

Nandana Mihindukulasooriya, MIT-IBM Watson AI Lab, Cambridge, USA

Adil Khadidos, King Abdulaziz University, Jeddah, Saudi Arabia

## Technical Program Committee

Sanju Tiwari, Universidad Autonoma de Tamaulipas, Mexico
Naziha Laaz, University Ibn Tofail, Morocco
Iurii Prokopchuk, Institute of Technical Mechanics of the National Academy of Sciences of Ukraine
Anil Nayak, University of Alberta, Canada
Kusum Yadav, University HAIL, Saudi Arabia
Filbert H. Juwono, Lecturer, Department of Electrical and Computer Engineering, Curtin University, Malaysia
Anatoliy Zabrovskiy, Petrozavodsk State University, Russia
Joel Luis Carbonera, UFRGS, Brazil
Nenad Petrović, Faculty of Electronic Engineering, University of Niš, Serbia

Alaa Khadidos, Information Systems Department, King Abdulaziz University, Jeddah, Saudi Arabia
Benjamin Warnke, Universität zu Luebeck, Germany
Ganapathy Mani, Purdue University, Indiana
Dac Nhuong, Haiphong University, Vietnam
Fatmana Şentürk, Pamukkale University, Turkey
Beakal Gizachew Assefa, School of Information Technology and Engineering, AAIT, Addis, Ababa, Ethiopia
Jyotirmoy Chatterjee, Lord Buddha Education Foundation (Asia Pacific University of Technology & Innovation), Kathmandu, Nepal
Ayodeji Salau, Afe Babalola University, Ado Ekiti, Nigeria
Bhagirathi Nayak, Sri Sri University, Cuttack, Odisha, India
Turupti Swarnkar, SOA, University, Odisha, India
Sirisha Potluri, Vignan Institute of Technology and Science, Hyderabad, India
Muhammad Imran Tariq, Superior University, Lahore, India
Asmita Yadav, JIIT, Noida, India
Monika Mangala, LToCE, Navi Mumbai, India
Sujata Chakrabrati, Centurion University, Odisha, India
Sachi Nandan Mohanty, Associate Professor, Vardhaman Engineering College, Hyderabad, India
Shajulin Benedict, Indian Institute of Information Technology, Kottayam, India
Prakash Gopalakrishnan, Amrita Vishwa Vidyapeetham, Tamil Nadu, India
Sachin Sharma, Manav Rachna International University, Bihar, India
Megha Bhushan, DIT University, Dehradun, India
Suja Radha, VIT University, India
Bunil Balabant Ray, NIT Meghalaya, India
Tanupriya Choudhury, UPES, Dehradun, India
Ravi Lourdusamy, Sacred Heart College (Autonomous), Tirupattur, Tamil Nadu, India
Shankey Garg, NIT, Raipur, India
Vijay Samule, NIT Bhopal, India
Dipak Mohanty, KIIT University, India
S. Shridevi, VIT, India
Gagandeep Singh Narula, GGSIPU, Delhi, India
Dr. Vijay Bhaskar Semwal, MANIT, Bhopal, India
Pawan Verma, GLA University, Mathura, Uttar Pradesh, India
Jatin Kumar Dash, SRM, Andhra Pradesh, India
Vinay Gautam, Chitkara University, Chandigarh, India
Ehtiram Khan, Jamia Hamdard University, Delhi, India
Chandra Yadav, JNU, Delhi, India
P. Shanmugaraja, Sona College of Technology, India
Amit Jain, Mandsaur Institute of Technology, Madhya Pradesh, India
Nafees Akhter Farooqui, Era University, Lucknow, India
Bhagirathi Nayak, Sri Sri University, Cuttack, Odisha, India
Nonita Sharma, NIT Jalandhar, India

Ruchi Mittal, Netaji Subhash Institute of Technology, Delhi, India
Dr. K. Kalai Selvi, Vels Institute of Science, Technology and Advanced Studies, Tamil Nadu, India
Nikita Mohod, Department of Computer Science and Engineering, Sipna College of Engineering and Technology, Amravati
Prajoy Podder, Institute of ICT, BUET, Dhaka, Bangladesh

# About This Book

The International Healthcare Informatics Conference (IHIC 2022)[1] is an annual international conference started this year looking at the increasing health concerns due to the COVID-19 pandemic. This is a sister conference of the International Semantic Intelligence Conference (ISIC 2022).[2] IHIC 2022 has been held during May 17–19, 2022, in hybrid mode at Sri Sri University, Cuttack, Odisha, India.

The COVID-19 pandemic has hit all the industry verticals ranging from banking, financial services and insurance (BFSI), government and defense, health care, manufacturing, retail, media and entertainment, transportation, energy and utilities, and others (travel and hospitality and oil and gas). All these listed organizations make use of information technology (IT) in one way or the other. The health industry also utilizes the complete tech-space in order to facilitate the required services. The International Healthcare Informatics Conference (IHIC 2022) offers the healthcare industry the means to stay ahead of the emerging threats to the well-being of humans. It provides a platform to the students, academicians, researchers, scientists, engineers, and industrial experts to share their knowledge and ideas in this field with the wider community for the benefit of mankind.

This book constitutes the proceedings of the 1st International Health Informatics Conference (IHIC 2022) held at Sri Sri University, Cuttack, Odisha, India from May 17 to 19, 2022. IHIC 2022 has four top-shot invited researchers as advisory. Being a sister conference of ISIC 2022, the participants of IHIC 2022 enjoyed the keynote talks and the tutorial sessions of ISIC 2022 as well. All in all, the conference witnessed six keynote speeches and four tutorials. All the speakers are multi-diversified in nature across the whole world and are esteemed experts in their field. The main conference organization has 13 chair members, whereas there are approximately 54 technical program committee members from various countries globally.

---

[1] https://www.ihic.semintelligence.org/.

[2] https://www.ifis.uni-luebeck.de/~groppe/isic/2022.

IHIC 2022 has been conducted in hybrid mode. Only high-quality manuscripts in the area of the conference are accepted for final publication by virtue of review and selection process. Every manuscript was reviewed by three to four reviewers with an acceptance rate of 50%. The volume comprises 78 manuscripts.

# Message from the Host Institution

The organizing committee is delighted to present the high-quality papers presented in the first Annual International Health Informatics Conference (IHIC 2022), a sister conference of International Semantic Intelligence Conference (ISIC 2022) being organized by Sri Sri University, Cuttack, Odisha, during May 17–19, 2022. The title was chosen based on its emergence and need of the hour in the current as well as next decade. Health informatics sector is considered as a pivotal area of research under the domain of computer science, social science as well clinical and public health practices. All these diversified fields need to be intertwined for different types of problem-solving purposes in order to cater to the needs of medical and healthcare practitioners. Clinical and public health informatics field has also seen a shift from a traditional silo approach to an integrated approach in collection, dissemination, and analysis of structured and unstructured information for overall health protection goals. Artificial intelligence, machine learning, and deep learning with its automated intelligent cognitive knowledge can be thought of as an assisting tool to the existing healthcare tools. The powerful technologies can be used to handle prearranged healthcare data, whereas unstructured healthcare datasets are managed by text mining and natural language processing. Also, rapid increase of healthcare data leading in the formation of outsized data is easily managed by big data and IoT-based methods.

IHIC aims to bring together researchers, practitioners, and industry specialists to discuss, advance, and shape the future of automated intelligent healthcare systems by virtue of AI, machine learning, deep learning, big data, and IoT. IHIC 2022 covers all dimensions of health informatics sciences in its three tracks, namely the research track, the trends and perspectives track, and application and deployment track. The research track covers areas such as AI, machine learning, deep learning, natural language processing, semantic web, cloud computing, Internet of thing, data mining and knowledge discovery, semantic and web technology, hybrid systems, agent computing, bio-informatics, and recommendation systems applications to healthcare sector. The trends and perspectives track explores the state of the art in the mentioned disciplines. The applications and deployment track showcases the latest advancements and applications of technologies in healthcare sectors with an intention that whenever any technology or methodology originates from the research community, its challenges and benefits are explored by its concrete usage in a practical setting and also sets the stage for its visibility. The “Call for Paper” for this conference was announced in the first week of September 2021, and we have kept a very deadline for paper submission, i.e., March 15. We have received 124 papers, which were considered for review and editing. Out of these 124 papers, 34 papers were accepted for the presentation and publication, which are covered in this proceeding. I am sure the participants must have shared a good amount of knowledge during the three days of this conference. I wish all success in their academic endeavor.

For Sri Sri University

Prof. (Mrs.) Rajita Kulkarni
President

# Keynote Talks

ISIC 2022 witnessed six keynote talks in six different topics.

a. **Title:** Applying Knowledge Graphs for Data Analytics and Machine Learning
   **Speaker:** Dr. Ernesto Jiménez-Ruiz (Lecturer, City, University of London)
   **Video Link:** **https://youtu.be/KSAKSOvCMHs**
b. **Title:** Detect, Characterize, and Accommodate Novelties in AI systems.
   **Speaker:** Bharat Bhargava (Purdue University, Indiana, United States)
   **Video Link:** **https://youtu.be/XWq8I-rv94U**
c. **Title:** Leveraging Artificial Intelligence and Machine Learning in Pandemics using COVID-19 as a Case Study
   **Speaker:** Sven Groppe (University of Lübeck, Germany)
   **Video Link:** **https://youtu.be/GAnC0ktljFU**
d. **Title:** Responsible AI for National Security
   **Speaker:** Amanda Muller (Artificial Intelligence Systems Engineer and Technical Fellow Northrop Grumman Mission Systems)
   **Video Link:** **https://youtu.be/PKPhCtG3EDo**
e. **Title:** Semantic Intelligence: The Next Step in AI
   **Speaker:** Sarika Jain (National Institute of Technology Kurukshetra, India)
   **Video Link:** **https://youtu.be/r18vXwkt57Y**
f. **Title:** Industrial Internet of Things and Digital Health Ecosystem: A Finnish Perspective
   **Speaker:** Rajeev Kanth (Savonia University of Applied Sciences, Kuopio Finland)

# Pre-conference Tutorials

IHIC 2022 witnessed four pre-conference tutorials.

1. **Tutorial 1: Building Domain-Specific Linked Data Applications**

**Presenters:** Sarika Jain, National Institute of Technology Kurukshetra, India; Pooja Harde, National Institute of Technology Kurukshetra, India; Ankush Bhist, University of Delhi, India; Nandana Mihindukulasooriya, MIT-IBM Watson AI Lab, Cambridge, USA
**Duration: 3 Hours**
**Video Link:** https://youtu.be/X9Y6snFFdUs, https://youtu.be/vpOADh4qh5E, https://youtu.be/8M5Xd9haF8k

2. **Tutorial 2: Knowledge Infused Reinforcement Learning for Social Good Applications**

**Presenters:** Manas Gaur, Research Scientist, Artificial Intelligence Institute, University of South Carolina; Kaushik Roy, Ph.D. Student at Artificial Intelligence Institute, University of South Carolina
**Duration: 2 Hours**
**Video Link:** https://youtu.be/A5SzrWBDxCY

3. **Tutorial 3: Knowledge Base Question Answering**

**Presenters:** Nandana Mihindukulasooriya, MIT-IBM Watson AI Lab, Cambridge, USA
**Duration: 1 Hour**
**Video Link:** https://youtu.be/2BknPDxaGUE

## 4. Tutorial 4: Modeling Epidemic Dynamics Thru Contact Networks

**Presenters:** Dr. Sharanjit Kaur, Acharya Narendra Dev College, University of Delhi, Delhi, India; Ms. Kirti Jain, Department of Computer Science, University of Delhi, India; Dr. Vasudha Bhatnagar, Department of Computer Science, University of Delhi, India
**Duration: 3 Hours**
**Video Link:** https://youtu.be/fwQccwh3EgU

# Contents

Development of Healthcare System Using Soft Computing Methods .... 1
Pallavi Mishra and Sachi Nandan Mohanty

A Novel Yoga-Based Practice Protocol to Quantify Stress After Performing Attention Task Using Non-invasive Technique ............ 13
Apurva Gupta, Virender Kadyan, and Shirley Telles

A Study of Deep Learning Algorithms in Sentiment Analysis of Diverse Domains ........................................ 23
Jutla Tanmaiyee, Ramya Sri Sushma Vasamsetti, Vangala Sai Pallavi, and Mukta Jagdish

Frequency Allocation in Cognitive Radio Networks .................. 31
Vemula Greeshma, Chennaboina Rahul, and Vemula Harini

An IOT-Based Advanced Health Monitoring System ................. 41
Ghaatki Vaishnavi, Ramavath Kavitha, and E. Vijaya Babu

Fault Tolerance in Vehicular Cloud Networks: Link Failure and Malicious Node Detection and Reconfiguration .................. 53
Sandeep N. Kugali, Bibhu Kalyan Mishra, Sneha Kadadevar, Piyush Kumar Pareek, and Malaya Malla

Multi-filter Enhanced Doctor and Patient Optimization Algorithm for Cancer Diagnosis ........................................ 69
Bibhuprasad Sahu, Saroja Kumar Rout, J. V. R. Ravindra, and Sachi Nandan Mohanty

Disease Diagnostic Model Using Fine-Tuned Intensive Learning for Medical Image ........................................ 79
Prasanalakshmi Balaji, Prabha Selvaraj, Vijay Kumar Burugari, and Kanmani Palaniappan

Challenges in Digital Health Care—A Hybrid-Decision Approach Understanding Key Challenges in India . . . . . 91
Sachi Nandan Mohanty, Tilottama Singh, Richa Goel, and Sukanta Kumar Baral

Telemedicine and Healthcare Setting for Remote Care of Patients: Status and Future Implications . . . . . 103
Sachi Nandan Mohanty, Venaik Anita, Sukanta Kumar Baral, Richa Goel, and Tilottama Singh

Performance Evaluation of ML Models in the Health Care Domain to Classify Liver Disease: A Case Study . . . . . 113
Goutam Datta and Anupam Singh

Telemedicine: Enhancing and Reimagining the Public Health in India . . . . . 121
Shweta Pandey, Abhishek Jha, Anjali Jha, Anurag Lokhande, Zahara Khan, Ulhas W. Karkhanis, and Aradhana Goutam

Effect of Point of Service on Health Department Student’s Creativity in Comprehensive Universities of Ethiopia: Moderating Role of Public-Private Partnership and Mediating Role of Work Place Learning . . . . . 135
Chala Wata Dereso, Shashi Kant, Maheswaran Muthuraman, and Gemechu Tufa

Prediction and Comparative Analysis of Few Most Impacted Countries by Coronavirus . . . . . 149
Simran Khanna, Meenal Garg, Seema Aggarwal, and Hardeo Kumar Thakur

An Efficient Method for Skin Cancer Detection Using Convolutional Neural Network . . . . . 161
C. Sreelakshmi, V. Adithya, and R. Rajesh

Collective Behavior in Community-Structured Network and Epidemic Dynamics . . . . . 167
Kirti Jain, Vasudha Bhatnagar, and Sharanjit Kaur

Incorporating Semantics for Text Classification in Biomedical Domain . . . . . 185
Shikha Verma and Aditi Sharan

COVID-19 Prediction from CT and X-Ray Scan Images: A Review . . . . 199
Prashant Kumar and Jagdeep Kaur

Design of a Prototypic Mental Health Ontology for Sentiment Analysis of Tweets . . . . . 215
Deepika Chaudhary, Jaiteg Singh, and Puneet Kaur

**An Ensemble for Attendance Management with Face Visualization and Recognition Using Local Binary Pattern Histogram** . . . . . . . . . . . . . . . 225
Guttula Nookaraju, T. Sreeja Reddy, D. Vignesh Kumar, and Madhu Bandari

**Healthcare Question–Answering System: Trends and Perspectives** . . . . . 239
Shrutikirti Singh and Seba Susan

**An Analysis of Word Sense Disambiguation (WSD)** . . . . . . . . . . . . . . . . . . . 251
Preethi Nanjundan and Eappen Zachariah Mathews

**Classification of Breast Invasive Ductal Carcinomas Using Histopathological Images Based on Deep Learning Techniques** . . . . . . . . 261
W. Jaisingh, N. Preethi, and S. Murali

**Heart Disease Detection Using Machine Learning Techniques** . . . . . . . . . 273
Satabdi Swain, Sujata Chakravarty, Bijay Paikaray, and Harshvardhan Bhoyar

**Evaluation of Spatiotemporal Fetal Cardiac Imaging Using Deep Learning Techniques** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 285
Dipak Kumar Nidhi, Khushboo Srivastav, Jukka Heikkonen, and Rajeev Kanth

**Application of Deep Learning on Skin Cancer Prediction** . . . . . . . . . . . . . . 299
Abhishek Dhiman and Naveen Chauhan

**Feature Analysis for Detection of Breast Cancer Thermograms Using Dimensionality Reduction Techniques** . . . . . . . . . . . . . . . . . . . . . . . . . 311
Vartika Mishra, Subhendu Rath, and Santanu Kumar Rath

**Covid-19 Question-Answering System Based on Semantic Similarity** . . . 323
Shrutikirti Singh and Seba Susan

**Impact of EEG Signals on Human Brain Before and After Meditation** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 331
Soumya Samarpita and Rabinarayan Satpathy

**COVID-19 Detection and Classification Method Based on Machine Learning and Image Processing** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 345
Inderpal Singh and Priyanka Chawla

**Optimal Convolutional Neural Network Model for Early Detection of Lung Cancer on CT Images** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 357
Dugyala Vignesh kumar, Guttula Nookaraju, T. Sreeja Reddy, B. S. Panigrahi, and S. N. Mohanty

**Lung Cancer Diagnosis Using Deep Convolutional Neural Network** . . . . 365
Jasdeep Kaur Basra, Eshaan Gupta, Munish Saini, and Amit Chhabra

**A Comprehensive Review of Brain Tumor Detection and Segmentation Techniques** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 381
Ankush Azade, Praful Kumar, and S. Sowmya Kamath

**Early Depression Detection Using Textual Cues from Social Data: A Research Agenda** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 393
Sumit Dalal, Sarika Jain, and Mayank Dave

# About the Editors

**Sarika Jain** graduated from Jawaharlal Nehru University (India) in 2001. Her doctorate, awarded in 2011, is in the field of knowledge representation in Artificial Intelligence. She has served in the field of education for over 21 years and is currently in service at the National Institute of Technology Kurukshetra (Institute of National Importance), India. Dr. Jain has authored or co-authored over 150 publications including authored and edited books. She has been a PI of projects sponsored by AICTE, DRDO, DST, and MHRD. Among the awards and honors, she has received are the Best Paper Award and the Best Faculty Award. She works in collaboration with various researchers across the globe and has been involved as a program and steering committee member at many prestigious conferences in India and abroad. Dr. Jain is a senior member of the IEEE, a member of ACM, and a Life Member of CSI.

**Sven Groppe** is a Professor at the University of Lübeck Germany. His publication record contains over 100 publications, including the book *Data Management and Query Processing in Semantic Web Databases* published by Springer. He was a member of the DAWG W3C Working Group, which developed SPARQL. He was the project leader of the DFG project LUPOSDATE and two research projects on FPGA acceleration of relational and Semantic Web databases and is a member of the Hardware Accelerator Research Program by Intel. He is currently the project leader of German Research Foundation projects on GPU accelerated database indices and the Semantic Internet of Things. Furthermore, he is leading a project about quantum computer accelerated database optimizations and he is a project partner in a project about COVID-19 high-quality knowledge graphs, visualizations, and analysis of the pandemic. He is also the chair of various workshops (SBD, VLIoT BiDEDE) at the ACM SIGMOD and VLDB conferences, and the general chair of the International Semantic Intelligence Conference (ISIC) in 2021 and 2022. His research interests include the Internet of Things, Semantic Web, query and rule processing and optimization, Big Data, Cloud Computing, peer-to-peer (P2P) networks, data visualization, and visual query languages.

**Nandana Mihindukulasooriya** is a research scientist at MIT-IBM Watson AI Research Lab, USA. Nandana holds a master's in Software Engineering and a Ph.D. in Artificial Intelligence. He has published over 60 peer-reviewed publications in highly reputed journals, conferences, and workshops related to topics such as Semantic Web and Knowledge Graphs, Neural-Symbolic AI, Knowledge Base Question Answering, Knowledge Induction from text, and information extraction. Before joining IBM, Nandana was a member of Ontology Engineering Group, Universidad Politécnica de Madrid (Spain) a leading European research group that is widely recognized in the areas of Ontology Engineering, Semantic Infrastructure, Linked Data, and Data Integration. He participated in several European research projects and Spanish national projects. Nandana has a keen interest in open-source and contributes to several open-source projects. Currently, he serves as a committee and a Project Management Committee (PMC) member of Apache Axis and Apache Web Services projects of Apache Software Foundation. He was also a speaker at ApacheCon US and ApacheCon Europe. Nandana is a senior member of the IEEE and a member of ACM and AAAI.

# Development of Healthcare System Using Soft Computing Methods

**Pallavi Mishra and Sachi Nandan Mohanty**

**Abstract** Real-world problems associated with healthcare domain are dynamic and non-deterministic in nature. Traditional analytics models are no longer effective for diagnostic processes and frequently require a large amount of computation time for decision-making, which cannot be sustainable in a technology-driven world. Soft computing methods, such as neural networks, machine learning, fuzzy logic, not only deal with the existing environment, but are also highly self-adaptive in terms of predicting and classifying new diseases that appear in the existing environment. As a result, these techniques are extremely beneficial and effective for use in the healthcare system. In this paper, we proposed a weighted averaging ensemble technique to perform supervised multi-level classification of dermatological diseases. Our experimental results show that the proposed method achieved significant better results than individual classifier models.

**Keywords** Erythemato-squamous diseases · Soft computing techniques · Neural networks · Ensemble classifiers

## 1 Introduction

Dermatology is the medical specialty that deals with skin, hair, and nails in their broadest sense. Skin disease has become a major issue for people all over the world [1]. Psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra pilaris are the most common erythemato-squamous diseases (ESDs) [2]. Most of these ESDs have similar clinical symptoms such as erythema (skin rashes) and scaling (dryness) in the early stages of the disease, but as the disease progresses, it begins to develop other clinical features, making it difficult

---

P. Mishra (✉)
Department of Computer Science and Engineering, Faculty of Science and Technology (FST), IcfaiTech, IFHE, Hyderabad, India
e-mail: Pallavi.mishra@ifheindia.org

S. N. Mohanty
School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, AP, India
e-mail: Sachinandan09@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_1

for dermatologists to determine which class of disease the patient has [3]. Therefore, it is a critical and challenging task for the medical experts (dermatologists) to distinguish them clinically in order to prescribe the proper treatment and medications. In many cases, steroids and antifungal medications are commonly used to treat various types of skin problems. This type of treatment is harmful to society since it increases the negative effects of local drugs like steroids, which are only effective at suppressing diseases rather than curing them permanently. As a result, this real-world challenge has been a focus of research, with many researchers aiming to construct sophisticated models that will aid medical experts in correctly classifying and diagnosing skin disorders based on symptoms for the benefit of the healthcare domain.

Until now, skin diseases were diagnosed so on the basis of the proficient doctors' long-term experience as well as the outcomes of clinical and pathological tests. Automatic detection or diagnosis of many diseases based on a list of symptoms can vastly improve and speed up the diagnostic procedure. Researchers have already conducted extensive research on this problem for the detection and recognition of various types of dermatological diseases.

Skin diseases have become a common ailment afflicting individuals of all ages. The early-stage identification and recognition of the symptoms would aid in proper diagnosis of the sickness would aid in the design of effective health management [4]. There has been a lot of interest in employing expert systems to conduct medical diagnosis and prescribe successful therapies in recent years. The development of a medical expert system as a knowledge base employing various soft computing approaches such as naïve Bayes [5], neural networks [6–8], fuzzy systems [3], genetic models promises to be a viable way for diagnosis and possible therapy.

With the evolution of soft computing techniques [9] and their successful application in healthcare field, the researchers put their efforts in developing a robust model using soft computing techniques to examine how these complexities of skin diseases diagnosis can be handled effectively. Soft computing techniques form a baseline for the development of computational intelligence (i.e., to build a sense of reasoning power) as like human brain. It has various components like neural network, support vector machines, fuzzy computing, genetic algorithms, etc. The unique feature of soft computing approach is its adaptability to uncertainty in the data and holds good for imprecise and partial truth and approximation input data. Due to the fact that the symptoms of skin diseases are uncertain and imprecise, therefore, the application of soft computing techniques would be proved significant in developing a model for classification [9].

Many research works based on skin disease classification have mainly implemented classifier models such as neural networks, support vector machines [4], nearest neighbor, naïve Bayes and achieved good accuracy. However, the combination of two or more classifier models (i.e., ensemble concept) has been proved effective over individual models in terms of performance.

In this paper, the proposed work is based on ensemble classifier which includes artificial neutral network, $K$-nearest neighbor [10], and support vector machine for skin disease classification. The dataset collected from UCI repository has 366

**Table 1** Class distribution of skin diseases

| Class code | Class of diseases | No. of samples |
|---|---|---|
| 1 | Psoriasis | 112 |
| 2 | Seborrheic dermatitis | 61 |
| 3 | Lichen planus | 72 |
| 4 | Pityriasis rosea | 49 |
| 5 | Chronic dermatitis | 52 |
| 6 | Pityriasis rubra pilaris | 20 |

instances with 6 target classes as shown in Table 1 [2, 3, 5] to prove insights regarding the number of patients suffering from different dermatological diseases. The dataset contains both clinical and biopsy, i.e., histopathological features as symptoms to classify the disease based on those symptoms. The challenging part lies in the fact that the earlier symptoms tend to vary and show symptoms of other type of skin diseases at subsequential stages, causing ambiguity for the experts in diagnosing the exact diseases at the early stage of the skin diseases.

The number of samples represents the total number of patients suffering from the corresponding class of diseases which again falls under either of 6 types of class codes.

## 2 Literature Review

Many studies on the diagnosis of erythemato-squamous diseases [3] have been conducted in depth by many researchers utilizing various models and approaches. The findings of the studies are summarized below.

The author [1] used bagging, AdaBoost, and gradient boosting classifier approaches to create an ensemble classifier using six different machine learning algorithms as separate classifiers for skin classification. When compared to individual classifiers, the model claims to achieve better results in ensemble mode.

The performance of CHAID decision trees [2] (DTs) for the analysis and diagnosis of ESDs is evaluated in this study. In terms of statistical accuracy, specificity, sensitivity, precision, true positive rate, true negative rate, and $F$-score, the prediction accuracies of these models are compared to the artificial neural network (ANN). Bagged ensemble surpasses other modeling methods in tests, according to the results.

A unique approach for the classification of erythemato-squamous diseases in an integrated fuzzy neural structure [3] (multi-input and multi-output) based on the Takagi-Sugeno-Kang (TSK)-type rule. The inference system's overall performance was empirically assessed in terms of classification accuracy, yielding a total accuracy of 98.37%. When compared to the results of other algorithms developed by other academics in the same domain, it was clear that the proposed algorithm was far superior.

An optimized support vector machine (SVM) was used to create an automatic erythemato-squamous illness identification system [4]. The swarm optimization algorithm, elephant herding optimization, was used to modify the parameters of support vector machines. When compared to the accuracy of other ways found in the literature, the prepared method excels and produces superior outcomes. Multi-layer feed-forward was used in this research. To categorize dermatology diseases, a neural network was deployed to a dermatology dataset retrieved from the UCI repository site. Working with a neural network that uses backpropagation for classification and prediction issues, algorithms offer the best results. It also has the ability to generalize and be applied to real-world issues.

On 366 occurrences of erythemato-squamous illnesses datasets, the study [5] employed naive Bayes, multi-layer perceptron, and J48 decision tree induction to create predictive data mining models. The baseline prediction performance of the classifiers was also evaluated using tenfold cross-validation and a set of performance criteria. The naive Bayes algorithm fared best, with a precision of 97.4%, followed by the multi-layer perceptron with a precision of 96.6%, and the J48 with a precision of 93.5%. Author [6] used soft computing techniques such as artificial neural networks and support vector machines to create models. The performance of ANN is more accurate for two hidden layers than SVM with RBF kernel, according to the results analysis.

To categorize dermatology illnesses, a multi-layer feed-forward neural network [7] was deployed to a dermatology dataset retrieved from the UCI repository site. For classification and prediction tasks, an artificial neural network using a back-propagation algorithm delivers the best results. The author [8] used ANN model for diagnosing skin illness and achieved 90% accuracy. Rambhajani et al. [9] has surveyed on various machine learning algorithms on dermatology dataset and by comapring past studies, the author concluded that nearest neighbor and naive bayes algorithms performed well with 99.25% accuracy. The paper [10] describes a unique approach for erythemato-squamous illness differential diagnosis. Fuzzy weighted preprocessing, nearest neighbor (KNN)-based weighted preprocessing, and a decision tree classifier are used in the proposed method. Using 20-fold cross-validation, the used decision tree classifier, fuzzy weighted pre-processing decision tree classifier, and KNN-based weighted preprocessing decision tree classifier achieved classification accuracies of 86.18, 97.57, and 99.00%, respectively. Vijaya [11] focuses on non-melanoma skin cancer and classifies types using support vector machine (SVM) to predict disease types effectively. Preprocessed skin lesion photographs are used to extract chrominance and texture attributes, and a training dataset is created.

Voting feature interval (VFFI5) is a new classification method created and applied to the problem of differential diagnosis of erythemato-squamous disorders. The domain comprises patient records with a known diagnosis [12]. The VFI5 classifier learns how to distinguish a novel case in the domain given a training collection of such records. VFI5 is a notion represented by feature intervals on each feature dimension individually. The VFI5 method uses real-valued voting to classify data.

Each feature participates in the voting process equally, and the predicted class is determined by the class that obtains the most votes. The VFI5 classifier's performance is empirically assessed in terms of classification accuracy and running time.

## 3 Dataset Description

The dataset used in this study is taken from UCI machine repository Güvenira and Demiröza [12]. This dataset is labeled (i.e., target values are present) and intended to classify various erythemato-squamous diseases [5]. There are 35 variables in this dataset, 34 of which are linear and 1 of which is nominal [1, 2]. Since the classes of the skin disease as mentioned in Table 1 are liable to share similar clinical features of scaling and erythema, thus identifying and diagnosing erythemato-squamous ailments can be challenging in dermatology. The six forms of skin illness include psoriasis, seborrhea, lichen planus, pityriasis rosea, chronic dermatitis, and pityriasis rubra [2]. One of the most prevalent therapies for these disorders is biopsy (i.e., tissue examination) [5]. In the early stages of a disease, a disease may have the characteristics of another class of disease, which is another challenge that dermatologists face when diagnosing these diseases. Patients were first analyzed using 12 clinical features, then 22 histological traits were assessed using skin disease samples [2]. The samples were examined under a microscope to identify histological features. If any diseases are discovered in the family, the family history attribute in the domain dataset is set to 1 (one), and if none are discovered, the value is set to 0 (zero). The patient's age is used to determine age characteristics [2]. Every additional clinical and histological trait was assigned a numerical value between 0 and 3 (0 = no features; 1, 2 = comparative middle values; 3 = highest value). The dataset of erythemato-squamous diseases has 366 occurrences and 35 attributes. Table 2 [3] summarizes the contents of the aforementioned attributes.

## 4 Proposed Work

The concepts of different classifiers used in the proposed work need to be explained to acquire insight knowledge of functioning of models (classifiers) for their efficient implementation in the classification of dermatological diseases.

### *4.1 Artificial Neural Networks*

Artificial neural network (ANN) [7] is a mathematical model that consists of number of artificial neurons interconnected each having associated with weights that can be

**Table 2** Summary of erythemato-squamous disease dataset

| Classes | Clinical attributes | Histopathological attributes |
|---|---|---|
| C1: psoriasis<br>C2: seborrheic dermatitis<br>C3: lichen planus<br>C4: pityriasis rosea<br>C5: chronic dermatitis<br>C6: pityriasis rubra pilaris | f1: erythema<br>f2: scaling<br>f3: definite borders<br>f4: itching<br>f5: Koebner phenomenon<br>f6: polygonal papules<br>f7: follicular papules<br>f8: oral mucosal<br>f9: knee and elbow<br>f10: scalp involvement<br>f11: family history<br>f18: hyperkeratosis<br>f34: age | 12: melanin incontinence<br>13: eosinophils in the infiltrate<br>14: PNL infiltrate<br>15: fibrosis of the papillary dermis<br>16: exocytosis<br>17: acanthosis<br>18: hyperkeratosis<br>19: parakeratosis<br>20: clubbing of the rete ridges<br>21: elongation of the rete ridges<br>22: thinning of the suprapapillary epidermis<br>23: spongiform pustule<br>24: munro microabcess<br>25: focal hypergranulosis<br>26: disappearance of the granular layer<br>27: vacuolization and damage of basal layer<br>28: spongiosis<br>29: saw-tooth appearance of retes<br>30: follicular horn plug<br>31: perifollicular parakeratosis<br>32: inflammatory mononuclear infiltrate<br>33: band-like infiltrate |

used for both classification and regression. The concept of neural network has been derived from the functioning and mechanism of human brain.

It functions with three layers, namely (i) input layer; that receives the input data, (ii) hidden layer; this layer is computational or processing layer, and (iii) output layer; it provides the final output [8] (Fig. 1).

The input layer transmits the information to the node of hidden layer and output layer, where actual processing takes place. The training phrase of the model starts by first assigning random values to weights for interlink between input variable to hidden node and hidden node to output node in feed-forward neural network. The output of each neuron is calculated by linear summation of inputs value and its corresponding weights and finally performing a linear or nonlinear activation function (i.e., sigmoid, tanh, ReLU, etc.) to control the output of model.

The activation functions on the output layer are chosen based on the nature of classification or prediction problem, i.e., either binary classification or multi-variate classification task. In case of neural networks, the architecture is drawn so as to apply activation functions on only hidden layer(s) and output layer (Fig. 2).

Different training algorithms are available such as Levenberg–Marquardt algorithm, which is one of the best-known algorithms in artificial neural network. The feed-forward neural network is used for training phrase where output of all nodes is calculated. Finally, output layer gives the overall output of the neural network.

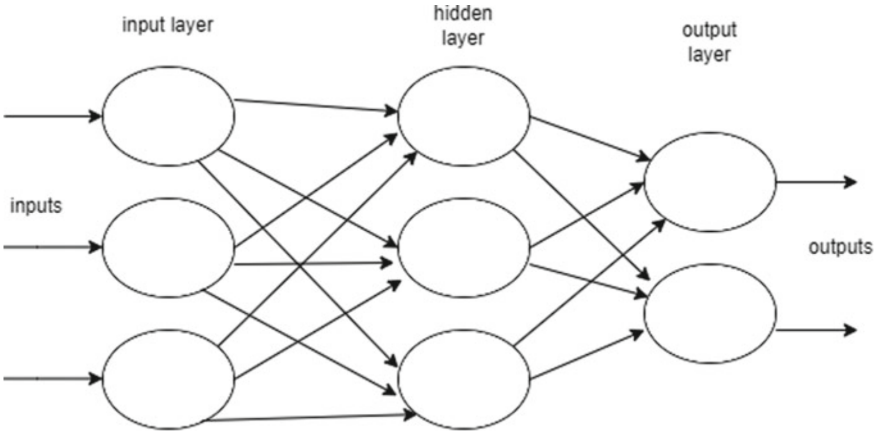
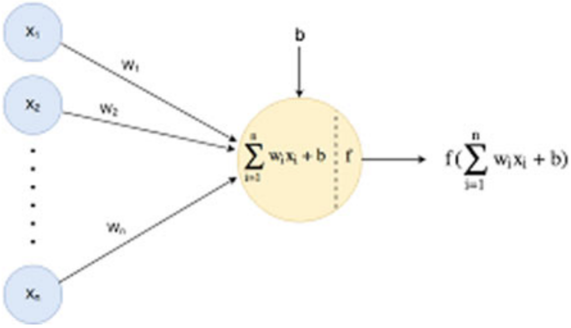


Fig. 1 Feed-forward neural network [6]

Fig. 2 Activation functions in neural networks [7]



Mathematically, the output is calculated as:

$$y = \text{activation function}\left(\sum_{i=0}^{n} x_i . w_i + \text{bias}\right) \tag{1}$$

where $x_i$: $i$th input value; $w_i$: weight of the $i$th input

The difference between the actual output and the obtained output is the error of the model. In order to minimize this error, gradient descent technique is used to update the weights till the error gets minimized. The error is calculated using minimum square error as loss function

$$\text{Error} = \min(\text{actual output} - \text{obtained output})^2 \tag{2}$$

The process of error minimization is possible using backpropagation approach. In this paper, error calculation is done using mean square error (MSE).

### 4.2 K-Nearest Neighbor

*K*-nearest neighbor (KNN) is another classifier model that does not require any explicit training for generalization of data points. It is one of the simplest classifiers that discovers the unidentified data points using previously known data points (i.e., nearest neighbor) and classifies them according to the voting system of their neighbors measured by a distance function. This distance function must minimize [10] the distance between the similar group of data points, while maximize the distance between dis-similar data points belonging to different classes. Many distance-based functions have been implemented such as Euclidean distance, Manhattan distance. The selection of the best distance function is one of the major factors that affect the KNN performance.

### 4.3 Support Vector Machine

Support vector machine is a supervised machine learning algorithm that supports both classification and regression analysis. It works well for binary classification where the hyperplane in a multi-dimensional space separates the different class levels. The decision factors for deciding the hyperplane are carried out by various kernel functions. The kernel functions play a significant role in transforming the input data into higher-dimensional form so that the decision boundary can be formed with the help of hyperplanes that classifies the data into distinct groups (Fig. 3).

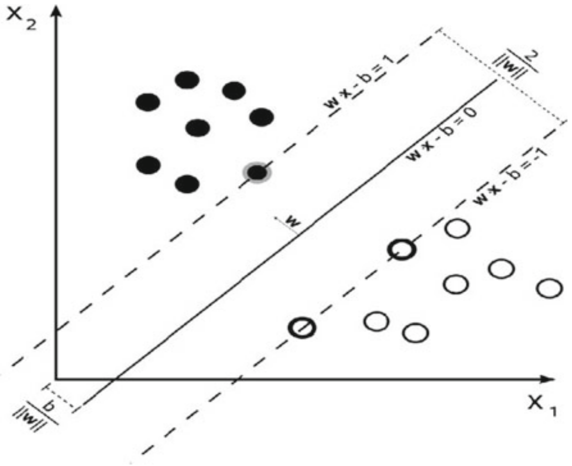


**Fig. 3** Maximum-margin hyperplane [6]

### 4.4 Proposed Model

Our proposed model is based on weighted averaging ensemble classifier technique. The ensemble classifier combines different classifiers such as artificial neural network (ANN), *K*-nearest neighbor (KNN), and support vector machine (SVM). The main objective of this proposed model is to improve the classification accuracy of the model using weighted averaging ensemble technique, where the higher weightage is assigned to the model whose performance is the best among other classifiers. Initially, the weight values (between 0 and 1) for the individual soft computing classifiers are assigned based on their accuracy value. The detailed working mechanism of the proposed model is explained using followed steps.

First step: Data normalization.

Second step: Splitting of dataset into training and testing data in the ratio 70:30.

Third step: Training phrase of three soft computing classifiers (ANN, KNN, SVM).

Fourth step: Evaluation of model on test data.

Fifth step: Accuracy obtained for each of the classifiers using formula.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

Sixth step: Once accuracy is known for each of the individual classifiers, then accordingly, weightage would be assigned.

Seventh step: Highest weigh value is assigned to the one having best accuracy, and second highest weightage is given to the model having accuracy lower as compared to the best one. Likewise, weights are assigned to each one of the individual classifiers.

Eighth step: Now, since ensemble model is the combination of all three classifiers, so the weighted average technique would provide with best result in terms of accuracy.

The schematic representation of the proposed model is shown Fig. 4.

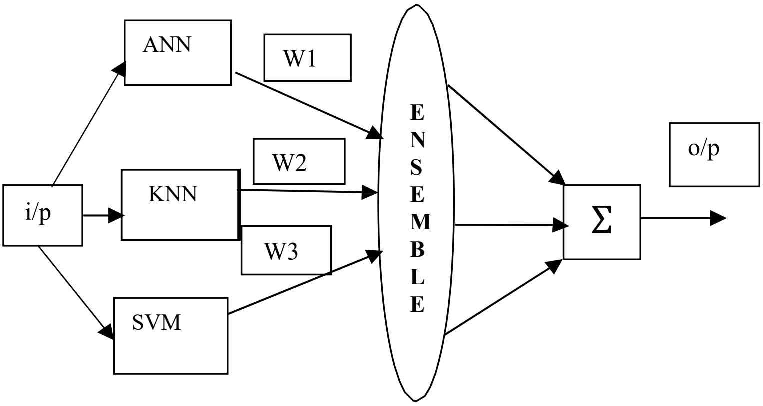


**Fig. 4** Proposed model

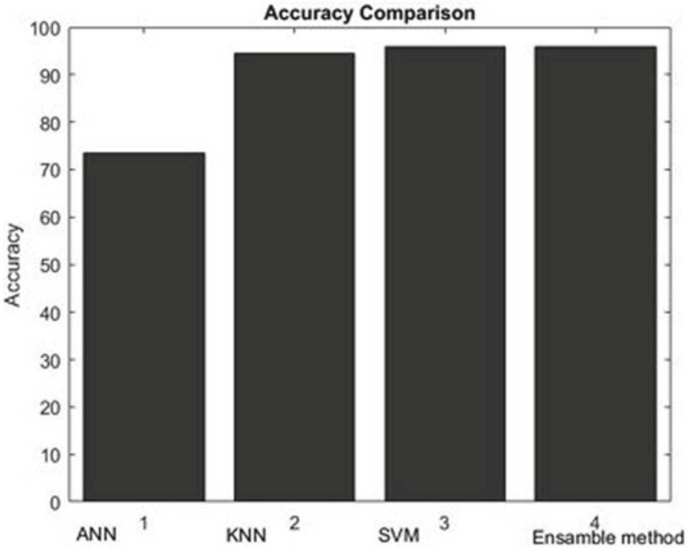


**Fig. 5** Comparison of accuracy of existing models and proposed model

## 5 Experimental Results

The proposed model described in Sect. 4.4 was implemented in MATLAB. The ESD dataset, which included 366 patient records and was retrieved from the UCI repository (erythemato-squamous dataset description, 2016), was used for training and testing. To find the best-suited classifiers, three individual classifiers were tested utilizing data samples. The ensemble's classifiers were trained and tested using tenfold cross-validation, with ninefold for training and onefold for testing. In terms of statistical correctness, the individual models' prediction accuracies are compared to the ensemble classifier. The weighted ensemble outperforms other modeling methods, according to the findings of experiments. The ANN achieved 73.43% accuracy, KNN outperformed ANN by achieving 94.5054%, and SVM achieved 96% accuracy.

The ensemble classifier (i.e., proposed model) performed outstandingly by achieving 96.3% accuracy, thus making the ensemble classifier model, the best model for classification. The results obtained by our proposed model (i.e., ensemble classifier) are compared against all three individual classifiers as shown in Fig. 5.

## 6 Conclusion

This paper proposed an ensemble classifier model based on artificial neural network (ANN), *K*-nearest neighbor (KNN), and support vector machine (SVM) with a view to achieve better accuracy in diagnostic process of dermatological diseases.

erythemato-squamous dataset [2] obtained from UCI repository includes all 34 attributes and 366 instances. Experiments were conducted using individual classifiers (ANN, KNN, SVM), and finally our proposed model (i.e., ensemble model) is experimented and compared with baseline models. The accuracy obtained by the ANN is 73.43%, while KNN improved the classification with accuracy 94.5%, and SVM outperformed with accuracy 96%. The overall accuracy obtained by the ensemble model is 96.3%. Thus, the type of dataset taken for the study determined the competence and performance of data mining techniques by using ensemble techniques. The accuracy of the classification can be further improved by using convolutional neural networks (CNN) and transfer learning technique to improve the accuracy of the result.

## References

1. Verma AK (2019) Classification of skin disease using ensemble data mining techniques. Asian Pacific J Cancer Prev APJCP 20.6
2. Elsayad AM (2018) Analysis and diagnosis of erythemato-squamous diseases using CHAID decision trees In: 2018 15th International Multi-Conference on Systems, Signals and Devices (SSD), IEEE
3. Idoko JB (2018) Fuzzy neural system application to differential diagnosis of erythemato-squamous diseases. Cyprus J Med Sci 3.2:90–98
4. Tuba (2017) Support vector machine optimized by elephant herding algorithm for erythemato-squamous diseases detection. Proc Comput Sci 122:916–923
5. Danjuma K (2015) Evaluation of predictive data mining algorithms in erythemato-squamous disease diagnosis. arXiv preprint arXiv: 1501.00607
6. Parikh (2015) Diagnosing common skin diseases using soft computing techniques. Int J Bio-Sci Bio-Technol 7.6:275–286
7. Patel (2017) An empirical study of applying artificial neural network for classification of dermatology disease. Indian J Sci Technol 10
8. Kabari LG (2009) Diagnosing skin diseases using an artificial neural network. In: 2009 2nd International conference on adaptive science & technology (ICAST), IEEE
9. Rambhajani M (2015) A survey on implementation of machine learning techniques for dermatology diseases classification. Int J Adv Eng Technol 8.2:194
10. Polat K (2006) The effect to diagnostic accuracy of decision tree classifier of fuzzy and k-NN based weighted pre-processing methods to diagnosis of erythemato-squamous diseases. Digit Signal Process 16.6:922–930
11. Immagulate I (2015) Categorization of non- melanoma skin lesion diseases using support vector machine and its variants. Int J Med Imag 3.2:34–40
12. Güvenira HA, Demiröza G (1998) Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals. Artif. Intell. Med. 13:147–165

# A Novel Yoga-Based Practice Protocol to Quantify Stress After Performing Attention Task Using Non-invasive Technique

**Apurva Gupta, Virender Kadyan, and Shirley Telles**

**Abstract** In today's era stress has become an integral part of everyone's life. Experiencing stress for longer period of time can create adverse effect on human physiology. To avoid its adverse effect, it is extremely important to understand physiology of stress. Also, it is very crucial to address all techniques which can help to deal with stress. Yoga is one of the scientific ways to reduce stress which has no side effect, but there is a lack of availability of properly detailed protocol. In this paper, a protocol is being detailed to quantify stress. This detailed protocol can guide researchers to design their own protocol which is very important for developing strong and result-oriented yoga-based study. This paper will discuss from yoga-based practices selection to participants selection to duration of practices to device used (parameters chosen) and some other minor but important parameters.

**Keywords** Yoga-based practices (YBPs) · Alternate nostril yoga breathing (ANYB) · Data analysis · Multi-modality · Single modality

## 1 Introduction

Stress is present in everyone's life on daily bases. Irrespective of age, it can be experienced from young children to old-aged people. Due to which current century is known as "stress era". Experiencing stress for shorter time period is beneficial as it can improve performance, prepare to fight, and sometimes create excitement. But if person experience any stressor very frequently and for longer period of time than this may lead to pathogenesis of chronic diseases like high blood pressure, depression, hypertension, diabetes, psychiatric diseases, or cancer [1].

---

A. Gupta (✉)
Machine Intelligence Research Center, School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
e-mail: apurva.gupta@ddn.upes.ac.in

V. Kadyan · S. Telles
Patanjali Research Center, Haridwar, Uttarakhand, India
e-mail: vkadyan@ddn.upes.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_2

Stress can occur due to any unwanted situation. These situations can be experienced from physical environment, due to mental instability or emotional conditions and could lead bodily or mental stress. Stresses can occur due to any external event (from the social situations, environment, or psychological) or internal events (from a medical procedure or illness). Stress is responsible to initiate the "fight or flight" response, a complex reaction of endocrinology systems and neurologic [2]. Based on the review of the literature elicited in [3], stress could be classified into three categories: 1—according to the nature of the stressor (psychological stress, physiological stress), 2—its impact on individual person (negative distress, positive eustress), and 3—according to the duration of stressor (chronic stress or long-term stress, acute or short-term stress).

There are *n* number of techniques available to deal with stress. Some of them are categories under clinical ways like taking some medicine, supplements, or going through with some clinical treatment, whereas some are categories under non-clinical ways to deal with stress like any form of physical exercise, good diet, having a strong social connection, doing meditation and yoga, being positive, and listening music [4].

## *1.1 Clinical Ways to Reduce Stress*

It includes taking pills, some physiological therapy, consulting to psychiatric, and it could be others as well. One of the clinical ways is to take GABA supplements. GABA refers to gamma-aminobutyric acid, and it is neurotransmitter or chemical messenger in brain which actually blocks some specific kind of signal in brain, which is responsible for anxiety, depression and provides a calming, resting state to mind [5]. People take GABA supplements to treat all disorders which are due to low GABA levels such as depression and anxiety [6].

## *1.2 Non-Clinical Ways to Reduce Stress*

### 1.2.1 Yoga-Based Practice

Yoga is an ancient Indian philosophy which is associated with three dimensions of individuals' body, mind, and soul to improve physical, mental, and spiritual health. Meaning of yoga is addition of all these three dimensions (body, mind, and soul) as it combines or works on all three dimensions. The practice of 6 yoga consists of a set of physical postures to maintain for specific period of time called "asanas", a voluntary control of breathing called "pranayama", a voluntary concentration of thoughts called "meditation", and repeated recital of phrases called "chanting/mantra" [7]. There are so many works done to know the efficacy of yoga to improve physiological parameters of person.

This paper is proposing yoga practices-based protocol which includes asana, breathing, and meditation. There are lots of evidences available to proof the efficacy of yoga in reducing stress. This protocol is designed to know the efficacy of yoga-based practices even after performing attention task. This work would able to tell that which yoga practice is able to minimize quantified stress after performing attention task**.**

## 2 Yoga and Stress Management

According to the ancient philosophy of India, yoga is an orthodox and non-invasive system of maintaining a healthier body that is aimed to "cut the seed of sorrow before it sprouts" to attain moksha. According to Srimad Bhagwat Geeta, yoga is defined as [8].

"yogah karmasu kaushalam", i.e., yoga is performing excellence in action.

"samatvam yogamuchyate" (Bh. Geeta 2/47), i.e., balanced state in all life situations like joy or sorrow.

"tam vidyat dukha sanyogaviyogam yogam sangyitam" (Bh. Geeta 6/23), i.e., yoga is detachment from miseries.

Despite its magnanimity in its core philosophy, twentieth-century population has emphasized the usage of yoga to promote health and as a tool for stress management along with the non-invasive approach to prevent and cure certain ailments. The philosophy of yoga emphasizes on an individuals' health condition with a primary focus on the present, by excluding past and future worries they have or may experience. Modern clinical sciences have explored this fact and made it evident when yoga is used as a means to reduce stress with the synergistic enhancement of complex human body structure. Yoga has been used not only to improve physiological and psychological parameters but has also been effective to manage stress, anxiety, depression, negative thoughts, to improve mental health, and to increase positive thoughts [9]. In recent years, yoga has become a tool for stress management not only in India but also in Western countries. The effectiveness of yoga to reduce stress has been proved clinically. Availability of clinical evidence proves the usage of yoga breathing to treat stress disorder, depression, and anxiety. In a study conducted by Shohani, 12 sessions of hatha yoga practice can reduce stress, depression, and anxiety significantly when given to 52 women 3 times per week [10]. In another study [11], 80 employees were categorized into two groups (yoga group and control group), and it was observed that 16 weeks of power yoga session (1 h, twice a week) can decrease stress, depression, anxiety, and insomnia significantly. In another study [12], mindfulness yoga intervention (yoga asana, pranayama, meditation, and mindfulness education) was carried out to address stress. Clinically, yoga has been proved to reduce physiological markers of stress which include heart rate, heart rate variability, cortisol, and systolic blood pressure by regulating the hypothalamic–pituitary–adrenal system and sympathetic nervous system [13].

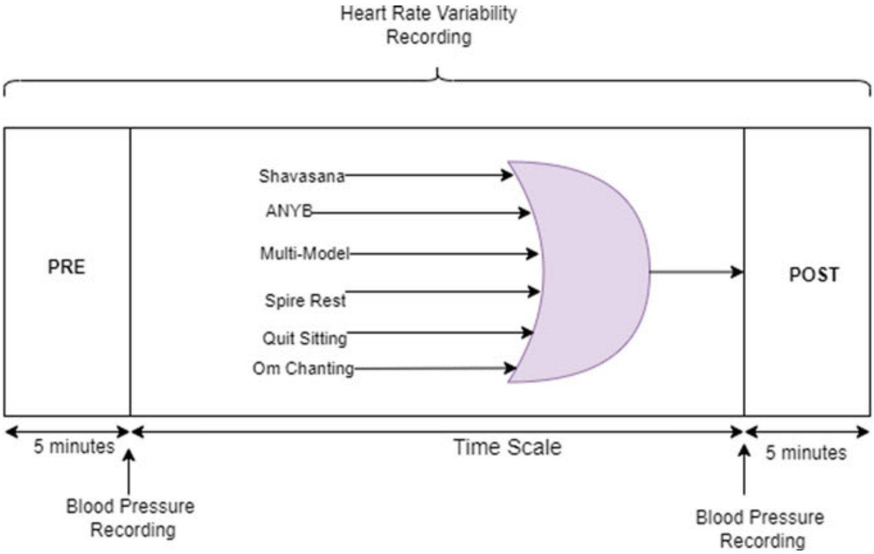


**Fig. 1** Basic building model of quantifying physiological and psychological stress using non-invasive technique

# 3 Methodology

## 3.1 *Study Design*

This protocol has been designed keeping multi-model framework in center. In this study, not only psychological data will be capture but also physiological parameters will also be captured. Also, some practices will be given as a single modality, and one practice will be given as multi-modality. A systematic representation of proposed protocol is defined in Fig. 1.

## 3.2 *Yoga Practices*

Protocol includes asana, pranayama, and meditation. In total, 6 interventions will be given as mentioned in Table 1. Protocol includes shavasana, ANYB, and Om chanting as single-modality yoga practice and MM as a multi-modality yoga practice which is a random combination of all these 3 yoga practices (shavasana, ANYB, and Om chanting). Spine rest and quite sitting are taken as a control sessions. Therefore in total subject will be given 6 interventions and one orientation session.

**Table 1** Protocol intervention list

| Intervention number | Intervention name | Intervention duration (min) |
|---|---|---|
| Intervention_1 | Shavasana | 15 |
| Intervention_2 | ANYB | 15 |
| Intervention_3 | MM | 15 |
| Intervention_4 | Spine rest | 15 |
| Intervention_5 | Quite sitting | 15 |
| Intervention_6 | Om chanting | 15 |

**Table 2** Inclusion and exclusion criteria

| | |
|---|---|
| Inclusion criteria | 28 < = Age < = 48<br>Familiar with basic yoga practices<br>Indian citizen<br>Willing to visit laboratory 7 times each for one hour for data collection<br>If female at 7th day after completing menstrual cycle |
| Exclusion criteria | Any psychiatric disorder<br>Performing yoga daily<br>Not willing to sign consent form<br>Any major surgery<br>Age < 28 and age > 48 |

## *3.3 Participants*

Participants are selected from the age range of 28–48 years. Proper inclusion and exclusion criteria were made which are mentioned in Table 2.

## *3.4 Data Collection*

To make it strong study, it is considering both physiological and psychological data. Under the category of psychological parameters to quantify stress, this protocol is using STAI, and under the category of physiological parameters, blood pressure and HRV are being recorded. In every session, data collection procedure will be done in the following given order from every subject:

a Pre-STAI collection
b Pre-recoding of HRV for 5 min
c Pre-recording of BP
d HRV recording during 15 min intervention.
e Post-HRV recording while performing attention task for 5 min
f Post-BP recording
g Post-STAI collection

### 3.5 Safety Consideration and Consent Form

Participants will be given all rights to withdraw if they would not feel comfortable at any stage of intervention. Along with consent form, participants will be given participants' information sheet where detailed description is given all about intervention like benefits of taking intervention, risk involved with study, all procedure from starting to end and other things. Also, confidentiality of identity of each participants will be maintained.

## 4 Discussion

### 4.1 Effect of YBP Using Physiological and Psychological Data

Evidence shows that yoga is effective in reducing stress, but some studies have been conducted only via considering physiological data (plasma cortisol, b-endorphin, interleukin (IL-6), tumor necrosis factor [TNF], blood pressure, heart rate variability, and others), whereas some studies have been conducted only via considering psychological data such as survey-scale PSS, STAI, Lovibond's stress scale.

To make this study strong, scientific, and justifiable, we aimed to collect both physiological and psychological data. Physiological parameters will be captured that keep most of the variables constant which can change the physiology of person.

### 4.2 Effect of YBP Using Single Modality and Multi-Modality Approach

Yoga practices are effective when given as a single modality or as a multi-modality. There is a lack of detailed description (practices, duration, and sequence) available when yoga intervention is given as a multi-modality. Also, there are number of parameters which are being considered to make protocol/intervention more string and effective like number of subjects, duration of intervention, length of duration, age range, control session, and others.

This protocol is strong due to several research points such as yoga intervention will be given as single modality and multi-modality; both physiological and psychological data will be captured; control session is being considered; and to avoid physiological variance, age range is kept closure, focusing to know the immediate effect of yoga intervention. Comparison is being presented in Table 3.

**Table 3** Comparison between existing protocol and proposed protocol

| Ref. | Both modality | Non-invasive | Subjects | Session duration (min) | Physiological and psychological parameters | Age range | Intervention duration | Control group |
|---|---|---|---|---|---|---|---|---|
| [14] | No | Yes | 42 | 10–40 | No | 23–67 | 12 weeks | No |
| [15] | No | Yes | 668 | 30–50 | No | 18–72 | 4–10 weeks | Yes |
| [16] | No | Yes | 226 | 140 | No | > 18 | 12 weeks | Yes |
| [17] | No | Yes | 191 | 60 | No | 18–22 | 12 weeks | No |
| Proposed | Yes | Yes | 60 | 15 | Yes | 28–48 | NA | Yes |

## 5 Conclusion

To strengthen the yoga-based research territory, it is important to build strong and well-defined protocol. This protocol is strong having multiple significant factors which includes single modality and multi-modality, significant number of subjects, combination of physiological and psychological parameters, short temporal scale intervention (this will help to know the immediate effect of yoga intervention), inexpensive, and non-invasive. To increase the practicality and adaptability of protocol, time scale of all intervention is kept 15 min. This protocol will help other researchers to build strong protocol. Also, this protocol can also be strengthened by doing some amendments which is kept for future research work.

**Acknowledgements** I would like to thank Mr. Sachin Sharma, Patanjali Research Foundation, for providing his valuable input during designing protocol.

## References

1. Nosaka M, Okamura H (2015) A single session of an integrated yoga program as a stress management tool for school employees: comparison of daily practice and nondaily practice of a yoga therapy program. J Altern Complement Med 21(7):444–449
2. shiel W (2002) Medicinenet. https://www.medicinenet.com/script/main/art.asp?articlekey=20104. Accessed 11 May 2020
3. Shahsavarani AM, Azad Marz Abadi E, Hakimi Kalkhoran M (2015) Stress: facts and theories through literature review. Int J Med Rev 2(2):230–241
4. Sharma DK (2018) Physiology of stress and its management. J Med Stud Res 1(001)
5. Nall R (2019) What to know about gaba. https://www.medicalnewstoday. com/articles/326847. Accessed 05 April 2020
6. Streeter CC, Jensen JE, Perlmutter RM, Cabral HJ, Tian H, Terhune DB, Ciraulo DA, Renshaw PF (2007) Yoga Asana sessions increase brain GABA levels: a pilot study. J Altern Complement Med 13(4):419–426
7. Rocha KKF, Ribeiro AM, Rocha KCF, Sousa MBC, Albuquerque FS, Ribeiro S, Silva RH (2012) Improvement in physiological and psychological parameters after 6 months of yoga practice. Conscious Cogn 21(2):843–850
8. Upadhyay D (2014) An approach to healthy life through yoga in ayurveda. Int J Res 2348–6848 Issn (online)
9. Duan-Porter W, Coeytaux RR, McDuffie JR, Goode AP, Sharma P, Mennella H, Nagi A, Williams JW (2016) Evidence map of yoga for depression, anxiety, and posttraumatic stress disorder. J Phys Act Health 13(3):281–288
10. Shohani M, Badfar G, Nasirkandy MP, Kaikhavani S, Rahmati S, Modmeli Y, Soleymani A, Azami M (2018) The effect of yoga on stress, anxiety, and depression in women. Int J Prev Med 9
11. Maddux RE, Daukantaité D, Tellhed U (2018) The effects of yoga on stress and psychological health among employees: an 8-and 16-week intervention study. Anxiety Stress Coping 31(2):121–134
12. Schuver KJ, Lewis BA (2016) Mindfulness-based yoga intervention for women with depression. Complement Ther Med 26:85–91
13. Pascoe MC, Thompson DR, Ski CF (2017) Yoga, mindfulness-based stress reduction and stress-related physiological measures: a meta-analysis. Psychoneuroendocrinology 86:152–168

14. Park CL, Finkelstein-Fox L, Sacco SJ, Braun TD, Lazar S (2021) How does yoga reduce stress? a clinical trial testing psychological mechanisms. Stress Health 37(1):116–126
15. Sahni PS, Singh K, Sharma N, Garg R (2021) Yoga an effective strategy for self-management of stress-related problems and wellbeing during COVID19 lockdown: a cross-sectional study. PLoS ONE 16(2):e0245214
16. Simon NM, Hofmann SG, Rosenfield D, Hoeppner SS, Hoge EA, Bui E, Khalsa SBS (2021) Efficacy of yoga vs cognitive behavioral therapy vs stress education for the treatment of generalized anxiety disorder: a randomized clinical trial. JAMA Psychiat 78(1):13–20
17. Tong J, Qi X, He Z, Chen S, Pedersen SJ, Cooley PD, Spencer-Rodgers J, He S, Zhu X (2021) The immediate and durable effects of yoga and physical fitness exercises on stress. J Am Coll Health 69(6):675–683

# A Study of Deep Learning Algorithms in Sentiment Analysis of Diverse Domains

**Jutla Tanmaiyee, Ramya Sri Sushma Vasamsetti, Vangala Sai Pallavi, and Mukta Jagdish**

**Abstract** Analysis of sentiment and opinion mining is the space of examination that investigates considerations, discernments, appraisals, practices, and feelings of individuals from the composed language on products, administration associations, human issues, occasions, points, and different traits. The main agenda of this paper is to depict the sentiments of people using deep learning algorithms and tools by classification methodology in categorizing polarity into positive, negative, and neutral.

**Keywords** Sentiment analysis · Machine learning · Social media

## 1 Introduction

Generally, sentiment analysis is the application of natural language processing and is also known as emotional extraction. Analysis of feeling can be achieved by the use of processes, specifically lexicon, and system gaining knowledge of Internet has end up ubiquitous in our lives these days [1]. Social network evaluation and text mining are used to conduct contextual modeling and knowledge extraction. Analyzing sentiments is to classify a given text into three lessons: goal, subjective, and polarity.

Analysis of emotions faces many limitations. The first challenge is a phrase of opinion taken into consideration could be fantastic in one case and dangerous in some other [2]. The second task is to explicit critiques in lots of approaches and to use traditional textual content processing to become aware of the difference among portions of textual content that do not exchange [3]. First, records are dynamically collected via web scraping, and then the polarity of opinion wishes to be analyzed.

---

J. Tanmaiyee · R. S. S. Vasamsetti · V. S. Pallavi
Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India

M. Jagdish (✉)
Department of Information Technology, Vardhaman College of Engineering, Hyderabad, India
e-mail: mukta.jagdish13@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_3

This paper provides a survey of sentiment evaluation on various domain names along with fitness merchandise, home equipment, eating places, movies, instructional gadget, political evaluations, e-commerce, food safety, and economic assessment [4].

## 2 Sentiment Analysis in Various Domains

### *2.1 E-Commerce*

It is not quick to Forecast sales of latest and present products, the usage of each numeric and textual data from customer evaluations. Types of algorithms used NLTK and SVM [5]. The NLTK corpus library includes sentences with positive and negative polarities scores that are used to train the version, that's then evaluate the usage of the SVM classifier.

### *2.2 Emotion Recognition in E-Learning Systems*

The approach of taking a pre-educated CNN, putting off in the long run linked layers, and using the outputs of CNN's penultimate layers as feature representation for images is known as switch mastering [6]. The deep CNN can still extract standard image functions like edges through freezing the weights of the convolutional layers. It might also help migrate classification records from an already skilled version on a specific dataset to the brand new dataset. During the improvement of facial features popularity system based on CNN, the system detected diverse feelings. The device changed into able to locate faces and classify emotions with greater accuracy [7].

### *2.3 Online Product Review*

Current approaches to online product reviews typically make empirical assumptions about the products of interest's sentiment characteristics, but they do not have enough detail about relative consumer preferences. Traditional machine learning techniques involve manual learning of the function, which takes more time and effort to predict product feedback [8]. CNN and the attention-based gated recurrent unit network are the algorithms involved.

## 2.4 Fake News Detection

Fake information determination and categorization based on medical posts on Twitter using just text. Personal experience is also needed to deduce the broader meaning of statements containing sarcasm, irony, and metaphor. One of the problems in identifying fake news on social media is labeling fake news manually, and misinformation disseminated by trusted news sources is difficult to identify as fake news. The algorithms used are LSTM, RNN, and CNN [9]. The approach of taking a pre-trained CNN, putting off in the long run linked layers, and the use of the outputs of CNN's penultimate layers as function illustration for photos is called transfer mastering [10].

## 2.5 Student Feedback

As the dimension of a phrase grows wider, traditional methods of assessing the relationship between words fail—the method before concentrates all of the phrases and produces the aim, that's a time-consuming operation [11]. The high-quality and accuracy of the results are solely structured on manual feature selection. The multi-head attention fusion model, glove an LSTM are the algorithms involved. Attention takes sentences, transforms them into a matrix with the words from one sentence forming the columns and the words from the other sentence forming the rows, after which suits them to discover suitable meaning. In machine translation, this is extraordinarily beneficial; it is better than before. The multi-head attention fusion model of word and glove and cove context embedding for LSTM is used. The multiple head attention mechanisms are used to capture the importance of each word and offer the importance of context words to predict the target automatically [12].

The rest of the application domains that depend on deep learning algorithms like CNN, LSTM, NLP, and so on are movie review, restaurant attribute classification, employee profiling framework, educational data, health-related tweets, financial review, multi-domain reviews, online product review, political review, hotel review, revealing fake news, sentiment classification with GST tweet data, estimate business sentiment from news texts, predict crisis related posts in social media, telemedicine demographic pattern, customer review, assessing the relationship between physicians and patients, medical comments, predicting the academic performance of students, emotion recognition in E-learning systems, prediction and analysis of financial time series, stock market prediction, public comments on hot topic, topic recognition, trip advisor, detecting personal intake of medicine from Twitter, predicting intensities of emotions and sentiments using stacked ensemble, personality detection, sarcasm classification, public opinion on work from home during COVID-19 pandemic, symbolic and sub-symbolic AI for product review, sentiment analysis on social media during COVID-19, food review, drug review, online travel product review, Amazon product review, and product review [13].

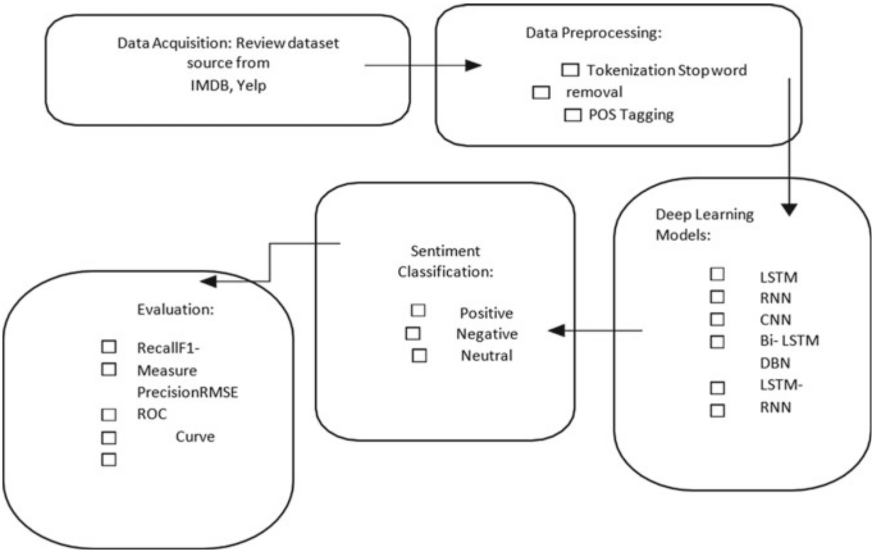


**Fig. 1** Methodology

## 3 Methodology

Methodology consists of data acquisition, data preprocessing, deep learning model, sentiment classifications, and evaluation (Fig. 1).

## 4 Discussion

Table four summarizes how the numerous classifier models used in numerous domain critiques had been evaluated. The accuracy of the E-commerce application product evaluation developed by means of classifying it with four unique metrics is 94%. In classification, consistency and accuracy are given by means of a big data collection and accuracy is 99.3%. The class model uses CNN and individual vector section pooling as data, and the end result accuracy is 88%. The classifier accuracy is expected the usage of a confusion matrix and the precise and wrong class of the full range of take a look at information instances, with a result accuracy of 78% (Fig. 2 and Table 1).

Precision, recall, $F1$-score, and accuracy were obtained using four steps in automated detection and sentiment class, yielding a result accuracy of 98%. A comparative analysis based on the accuracy of different models is shown in figure. On small datasets, LSTM outperforms CNN on health-related datasets. CNN s have proved to be an effective solution; however, they require huge records due to many parameters. As an end result, CNN does now not reach excessive precision inside the case of

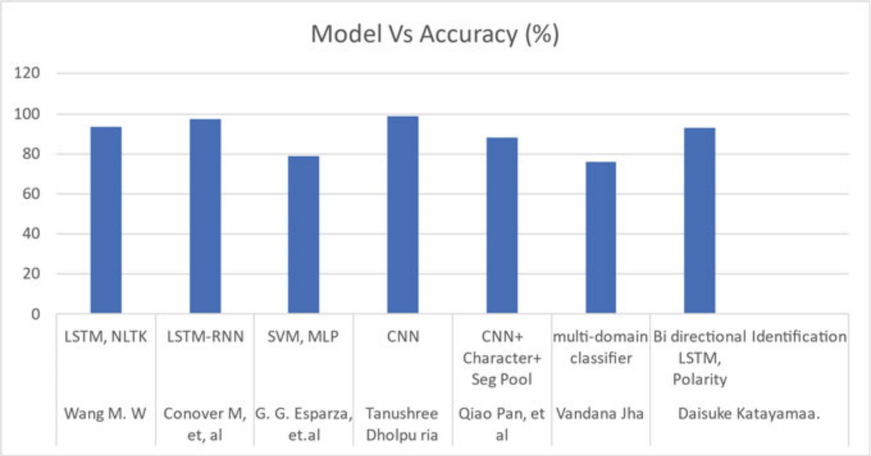


**Fig. 2** Model vs accuracy

**Table 1** Deep learning models with accuracy

| Author | Deep learning models | Accuracy (%) |
|---|---|---|
| Wang | LSTM, NLTK | 93.5 |
| Conover et al. | LSTM-RNN | 97.4 |
| Esparza et al. | SVM, MLP | 79 |
| Tanushree Dholpu ria | CNN | 99 |
| Qiao Pan et al. | CNN + Character + Seg pool | 88 |
| Vandana Jha | Multi-domain classifier | 76 |
| Katayamaa | Bidirectional LSTM, polarity Identification | 93 |

small datasets. Since small datasets are often vulnerable to the specifics of training information and random initialization, overfitting is a significant problem. It is not easy to see how CNN can only function well with vast data to ensure that the proposed model is robust. If the dataset is unbalanced, CNN compares outcomes first with two positive and non-positive classes, positive, non-positive, and neutral. In LSTM, there are both positive and negative effects (Figs. 3 and 4; Tables 2 and 3).

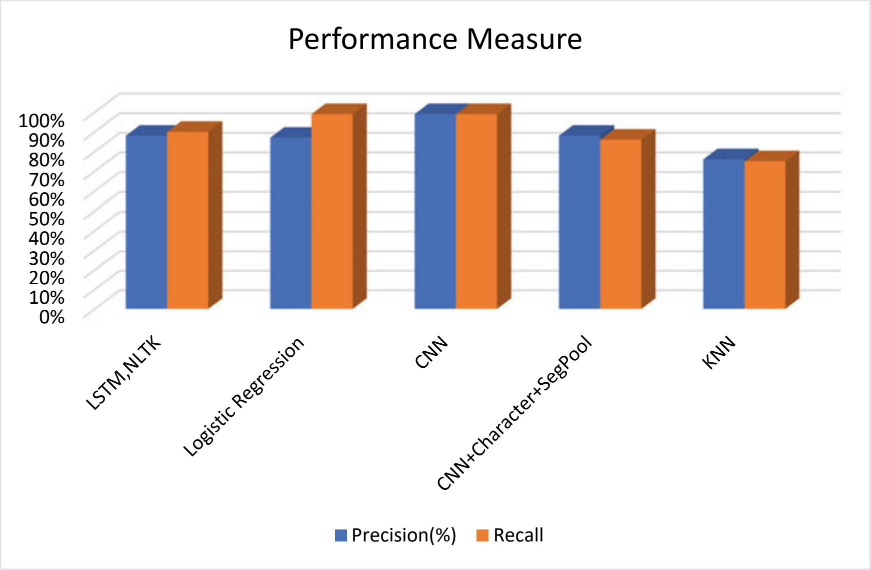


**Fig. 3** Performance measure

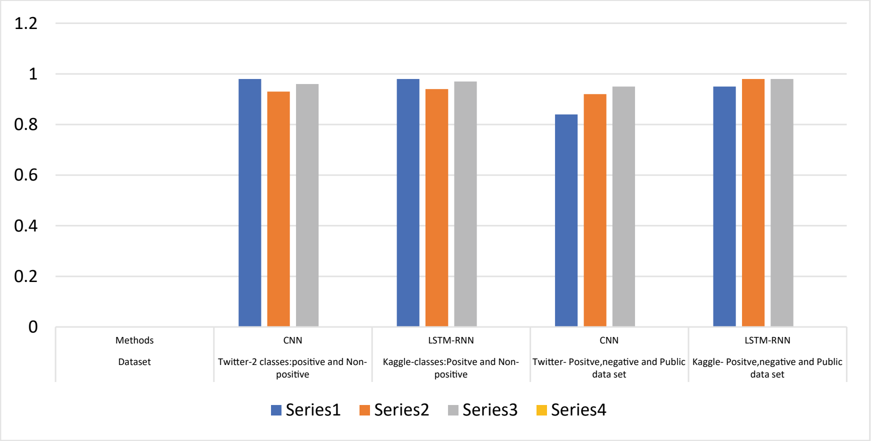


**Fig. 4** Overall analysis

**Table 2** Classification, precision, recall

| Classifier | Precision (%) | Recall (%) |
|---|---|---|
| LSTM, NLTK | 88 | 90 |
| Logistic regression | 87 | 99 |
| CNN | 99 | 99 |
| CNN + Character + Seg pool | 88 | 86 |
| KNN | 76 | 75 |

**Table 3** Dataset, methods, precision, recall, and accuracy

| Dataset | Methods | Precision | Recall | Accuracy |
|---|---|---|---|---|
| Twitter-2<br>Classes: Positive and non-positive | CNN | 0.98 | 0.93 | 0.96 |
| Kaggle classes: Positive and non-positive | LSTM-RNN | 0.98 | 0.94 | 0.97 |
| Twitter: Positive, negative, and public dataset | CNN | 0.84 | 0.92 | 0.95 |
| Kaggle: Positive, negative, and public dataset | LSTM-RNN | 0.95 | 0.98 | 0.98 |

## 5 Conclusion

A comparative study of the deep learning models associated with the techniques of meaning classification is provided in this paper. Deep learning concepts have achieved a significant perception in the research areas using multiple classification methods and techniques. This paper comparatively assesses different classifications and deep learning models by analyzing method performance using metric parameters such as precision, recall, and accuracy. With downloaded sentiment prediction datasets from platforms such as Amazon and Yelp, many algorithms for analyzing feeling and classification are introduced. The paper also includes a survey on different models implemented from 2015 to 2019, where a different set of algorithms were proposed in each paper. Deep learning models like autoencoders, deep belief networks, CNN, RNN, and classifiers like SVM, Naive Bayes, logistic regression, KNN, ANN, MLP, etc., are studied. Observations of possible solutions are seen in the tables provided, including the advantages and disadvantages. The classifiers and algorithms specified that deep learning models have important implications in the analysis of sentiments.

## References

1. Hassan A, Mamhmood A (2018) Convolutional recurrent deeplearning model for sentence classification. IEEE Access 6:13949–13957. https://doi.org/10.1109/ACCESS.2018.2814818
2. Valdivia A, Luzón MV, Herrera F (2017) Sentiment analysis in tripadvisor, affective computing and sentiment analysis. IEEE Intell Syst 72–77
3. Shin B, Lee T, Choi JD (2017) Lexicon integrated CNN models with attention for sentiment analysis. 149–158
4. Baccouche A, Garcia-Zapirain B, Elmaghraphy A (2018) Annotation technique for health-related tweets sentiment analysis 1978-5386-75687/18. https://doi.org/10.1109/ISSPIT.2018.8642685
5. Soh C, Yu S, Narayanan A, Duraisamy S, Chen L (2019) Employee profiling via aspect-based sentiment and network for insider threats detection. Exp Syst Appl 1–15. https://doi.org/10.1016/j.eswa.2019.05.043
6. Varshneya D, Dhananjay P, Jayagopi DB (2016) Restaurant attribute classification using deep learning. 978-1-5090-3646-2. https://doi.org/10.1109/INDICON.2016.7839050
7. Katayamaa D, Kino Y, Tsuda K (2019) A method sentiment polarity identification in financial news using deep learning. Proc Comput Sci 159(2):1287–1294

8. Mahata D, Friedrichs J, Shah RR, Jiang J (2018) Detecting personal intake of medicine from twitter, department: affective computing and sentiment analysis. IEEE Intell Sys 87–95
9. Cambria E, Li Y, Xing FZ (2020) SenticNet 6: ensemble application of symbolic and subsymbolic AI for sentiment analysis 105–114
10. Esparza GG, Díaz AP, Canul-Reich J, De-Luna CA, Ponce J (2016) Proposal of a sentiment analysis model in tweets for improvement of the teaching- learning process in the classroom using a corpus of subjectivity. Int J Comb Optim Prob Inf 7(2):22–34, ISSN: 2007-1558
11. Amrieh EA, Hamtini T, Aljarah I (2016) Mining educational data to predict student's academic performance using ensemble methods. Int J Database Theor Appl 9(8):119–136
12. Talpada H, Halgamuge MN, Vinh NTQ (2019) An analysis on use of deep learning and lexical-sentiment based sentiment analysis method on twitter data to understand the demographic trend of telemedicine
13. Jagdish M, Shah DU, Agarwal V, Loganathan GB, Alqahtani A, Rahin SA (2022) Identification of end-user economical relationship graph using lightweight blockchain-based BERT model. Comput Intell Neurosci Hindawi 2022:9. Article ID 6546913

# Frequency Allocation in Cognitive Radio Networks

**Vemula Greeshma, Chennaboina Rahul, and Vemula Harini**

**Abstract** Frequency sensing is a dynamic exploration region in cognitive radio organizations that begins the compelling information splitting between authorized and unlicensed cognitive network clients. Increasing the likelihood of identifying phony alerts is a hasty test of a major chunk of the heinous detecting strategies. The research offers the KHWO-ACNN-based organization as a spectrum detecting approach. The abandoned range is not fixed in stone using the proposed technique, which promptly allocates the free range groups to the key clients with the purpose of limiting the postponement due to the effective working of the combination place. The Eigen-esteem-based helpful detecting is executed in the intellectual radio for compelling detecting. The examination of the proposed technique is progressed based on presenting measurements such as bogus alert rate probability and detection of likelihood probability. The proposed spectrum detecting strategy out-flanks the current strategies that yield a most extreme probability of likelihood and least bogus alert rate probability as 0.9805 and of 0.009.

**Keywords** Frequency sensing · Cognitive radio networks

## 1 Introduction

Cognitive radio is a climate which comprehends the environmental elements and knows about its own current circumstance. Intellectual radio is utilized to limit the impedance which improves the range use by unique re-arrangement. In the cognitive radio networks, there are two sorts of buyers. They are the authorized users and unauthorized users. Cognitive radio is an innovation which allows the unauthorized users who do not have the permit to get to the range of the authorized users. The unauthorized users in cognitive radio networks sub-partition the channels of authorized users

---

Chennaboina Rahul, Vemula Harini: These authors contributed equally

V. Greeshma (✉) · C. Rahul · V. Harini
Electronics and Communication Engineering, Vardhaman College of Engineering, Kacharam, Hyderabad, Telangana 501218, India
e-mail: 1002.greeshma@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_4

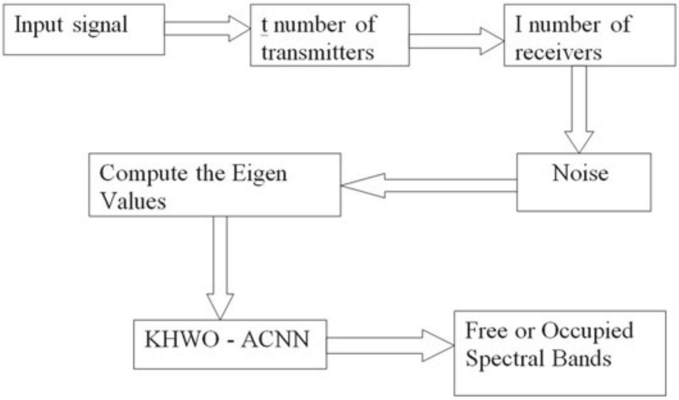


**Fig. 1** Block diagram

as sensible openings. The unauthorized users detect the exercises of authorized users in these openings for a brief time frame and therefore, access those openings when inactive that are named as range opening of void areas. Accordingly, ensures that the unauthorized users access the authorized groups without intruding on the authorized users. In light of the strategies of the range versatility, cognitive radio users pull out their individual channel, when the essential clients come back, guaranteeing the need of essential clients and also, it guarantees these assurance of essential client's transmission. Cognitive radio utilizes a spectrum detecting technique, known as unique range access strategy, to deal with the accessibility of the range and thus, to limit the issues because of ghastly blockage. Spectrum detecting is the critical capacity in cognitive radio organizations to such an extent that it empowers the unauthorized users to track down the unused range groups with permit and accordingly, to shield the primary user's transmission (Fig. 1).

## 2 Results

In order to increase the possibility of recognition and reduce fictitious caution, we established the idea of spectrum detecting in cognitive radio organizations. Utilizing the krill herd whale improvement computation, this probability of likelihood and probability of bogus warning rate are separated. Based on our findings, we can conclude that the krill herd whale improvement computation provides us with the real upsides of probability of likelihood and probability of false alert rate when compared to other strategies like GLRT, MMED, and MED. Here, we put up a channel model and use a method to carry out foreign detection in the CRNs. The channel model is dependent on the transmitter and information signal used for channel correspondence. In order to ensure that the spectrum sensing is ready for correspondence, phantom detection works to prevent delay and correspondence cross-over. Using the FC, which produces information on the ranges the PUs cover, it is possible to detect the presence of free spectral groups. The FC gathers information about the SUs' occupancy of the range from them. Finally, the optimal tuning of the vacant groups is carried out once the vacant groups are discovered using the suggested KHWO-based ACNN. The effectiveness of the built-in approach for extraterrestrial detecting is guaranteed by the analysis of recognition likelihood and bogus alert likelihood. For the suggested technique, the chance is high, but the risk of a false alarm rate is low. Regarding the gamma's worth of 0.85, the proposed technique's possibility of likelihood and possibility of bogus alert rate are found to be 0.9805 and 0.009, respectively.

### 2.1 Probability of Likelihood

The ratio of the whole number of identified ranges to the total number of freely accessible ranges in the CR climate is known as the possibility of likelihood of the accessible range. The possibility of likelihood metric needs to provide strong motivation for the expansion of the display.

$$\text{Probability of likelihood} == \frac{\text{Total number of identified spectrums}}{\text{Total number of unused spectrums}} \tag{1}$$

### 2.2 Probability of Bogus Alert Rate

Better execution is evidenced by the chance of false alert rate's more modest value. The proportion of the total number of ranges incorrectly identified as the accessible range to the absolute groups in the range is known as the false alert rate in the recommended strategies.

$$\text{Probability of bogus alert rate} == \frac{\text{Number of falsely identified spectrums}}{\text{Total number of unused bands}} \tag{2}$$

## 3 Explanation

The proposed model of the cognitive radio framework utilizes the unified helpful framework detecting that is portrayed and the idea of cognitive radio is to deal with the radio range lack through the distinguishing proof of new ways for getting to the unused frequency bands. The cognitive radio organizations comprise of the authorized users and the unauthorized users alongside their individual transmitters and a concentrated fusion center [1]. The best technique for identifying the presence of the involved or empty range depends in the location of the authorized users [2]. The unauthorized users notice the authorized users who will use the specific band prompting the discovery of the involved or empty frequency bands [3]. The fundamental thought process of the spectrum detecting is to decide the empty frequency bands for the correspondence and allot these empty frequency bands ideally with no impedance to the authorized users [4]. At first, the authorized users impart their utilization data to the primary user transmitter, which communicates the data to the fusion facility [5]. Thus, the fusion facility centers collect the request information from unauthorized users, checks the frequency bands used by authorized parties, and reports the information regarding the availability of the spectral groups [5]. The accessibility of the empty frequency bands depends on the energy-based spectrum detecting [6]. When a client requests a band for communication, the cluster really examines the availability of the empty spectral band or collects the information from authorized parties and reports it to the vacant spectral band to ensure that there is no cross-over or delay in communication [7]. The energy regulates what spectrum ranges or networks are accessible [8]. The information signal is given to the '$T$' number of transmitters, which is then given to the '$I$' number of receivers. From here, the data is given to the fusion center. The commotion, Eigen esteems, KHWO calculation, and ACNN are processed to the fusion center.

## *3.1 Actor Critic Neural Network (ACNN)*

The discrete-time memory-less multi-input multi-output model is used to calculate the energy measurements. The sensors are used to collect data from users and send it via transmitters along the transmission line. The yield from the recipients is combined with the additional material turmoil and delivered to the ACNN block, which does the dreadful detection.

## *3.2 Actor*

The presenter employed a specific method known as a neural network, and the actor-critical network slots in as an operator module using the neural organisation with the single convolutional. The operator component's contribution conveys the data from the test insights and the Eigen benefits of the network scenario. Along these lines, the input and concealed layer's progressive calculations add to the yield, which forms the contribution to the critic layer.

### *3.3 Critic*

This layer appreciates the actor layer's yield as well as the actor layer's contributions to foreseeing the free unearthly band. The variable loads are used to run a smooth prediction with greater precision.

### *3.4 Krill Herd Optimization*

This optimization depends on crowding conduct of the krill crowd that depends on the natural and ecological cycles. The goal of the krill herd calculation is to increase krill thickness while searching for resources that are calibrated to global optimum. Every krill proceeds in this manner while the interaction of the inquiry, keeping in mind the greater thickness and global optimum. The target work's base value suggests a high thickness krill group and supply are nearby to the krill. The development cycle is illustrated as follows: In the first step, the krill population is established, and the krill group takes shape. Poaching is indeed a cycle that reduces krill thickness and ensures less space between the krill and the food source. The distance between the krill and where the food should be at all times determines the fitness in the krill herd calculation, where it should be less. The development of the krill is dependent upon three key factors, such as rummaging behavior, irregular dispersal, and development dependent on the other individual krill.

### *3.5 Whale Optimization*

The mound-back whales' foraging actions can affect whale optimization. Gray whales can identify their predator's location and enclose it. The placement of arbitrary configurations serves as the foundation for whale optimization. The whale optimization takes the current arrangement as the objective arrangement and moves forward on the interaction dependent on the present arrangement because the initial area of optimal arrangement is unknown. At every emphasis, the situation of the pursuit specialist is refreshed by either arbitrarily picked specialist or the best arrangement. This interaction is rehashed until the best arrangement comes to. The whale optimization calculation is more aggressive when contrasted and different techniques and it can tackle the additional difficult issues. The whale optimization calculation benefits from a decent equilibrium between investigation and double-dealing, which assists this calculation with tracking down the worldwide optimum. Whales are found in gatherings and show the uncommon hunting conduct named as the scrounging conduct called bubble-net taking care of strategy.

### 3.6 KHWO-ACNN

This calculation hinders assembly to the worldwide ideal and it targets processing the ideal loads so that it could come into actor and critic layer of ACNN. There are four stages in the KHWO-ACNN calculation. They are

Initial population: The instatement will be initial procedure which introduces size of the population which alludes to loads of ACNN. Fitness Evaluation: A wellbeing metric is determined by the base value of error, which is determined by the yield from the pundit/entertainer layers as well as the ground truth. The wellness metric should be a starting point.

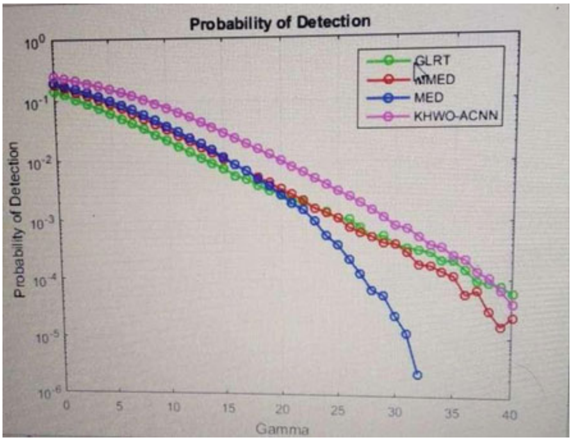
Updating the weights: The weights are updated by utilizing the proposed calculation. Whale enhancement is utilized for the position update and is changes as per the krill herd calculation. The proposed position update improves the union of global optima to the point that the ACNN's versatile loads can be deduced.

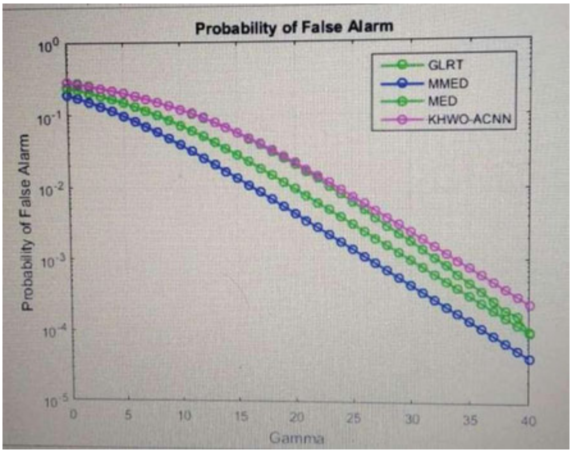
Completion: This interaction is repeated as many times as possible until the global ideal arrangements are discovered.

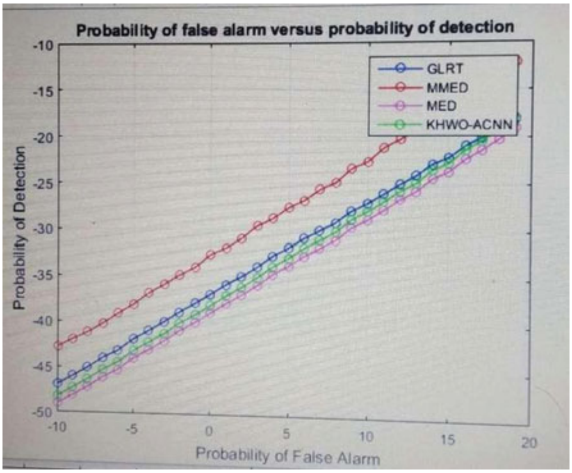
## 4 Conclusions

### 4.1 Analysis Using 20 Users

For 20 users, the probability of bogus alert rate is 5% and the probability of likelihood is 0.98 for GLRT, 0.8 for MMED, 0.9782 for MED and 1 in this technique.

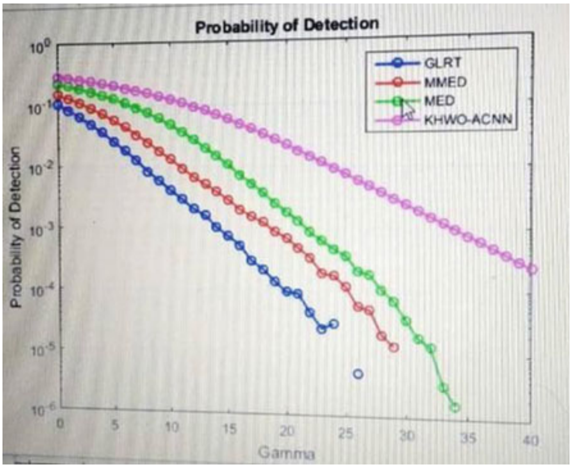


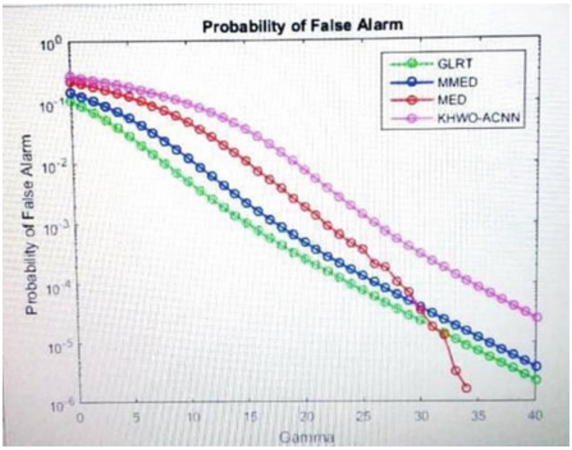


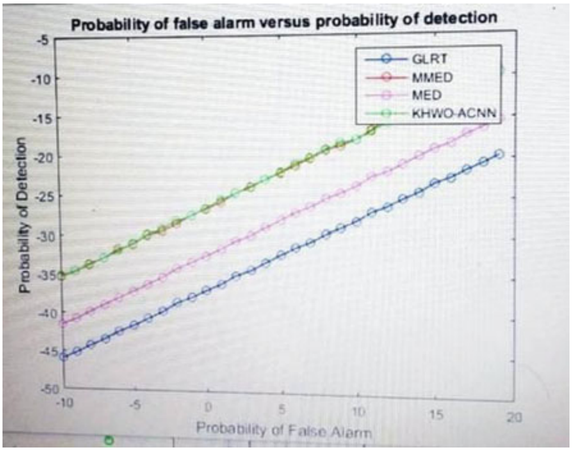


## 4.2 Analysis Using 40 Users

For 40 users, the probability of bogus alert rate is 5% and the probability of likelihood is 0.9489 for GLRT, 0.5757 for MMED, 0.7816 for MED, and 0.99 in this technique.


Probability of Detection
Probability of Detection
Gamma
GLRT
MMED
MED
KHWO-ACNN
$10^0$
$10^{-1}$
$10^{-2}$
$10^{-3}$
$10^{-4}$
$10^{-5}$
$10^{-6}$
0
5
10
15
20
25
30
35
40


Probability of False Alarm
Probability of False Alarm
Gamma
GLRT
MMED
MED
KHWO-ACNN
$10^0$
$10^{-1}$
$10^{-2}$
$10^{-3}$
$10^{-4}$
$10^{-5}$
$10^{-6}$
0
5
10
15
20
25
30
35
40



# References

1. Pandi N, Kumar A (2017) A review on cognitive radio for next generation cellular network and its challenges. Am J Eng Appl Sci 10(2):334–347
2. Marinho J, Granjal J, Monteiro E (2015) A survey on security attacks and countermeasures with primary user detection in cognitive radio networks. EURASIP J Inf Secur 2015(1):1–14
3. Zhang D, Chen Z, Ren J, Zhang N, Awad MK, Zhou H, Shen XS (2016) Energy-harvesting-aided spectrum sensing and data transmission in heterogeneous cognitive radio sensor network. IEEE Trans Veh Technol 66(1):831–843
4. Supraja P, Jayashri S (2017) Optimized neural network for spectrum prediction scheme in cognitive radio. Wirel Pers Commun 94(4):2597–2611
5. Roy P, Muralidhar M (2015) Channel state prediction in a cognitive radio network using neural network levenberg-marquardt algorithm. Int J Wirel Commun Netw Technol 4(2):24–29
6. Gavrilovska L, Atanasovski V (2011) Spectrum sensing framework for cognitive radio networks. Wirel Pers Commun 59(3):447–469
7. Ivrigh SS, Sadough SM-S (2012) Spectrum sensing for cognitive radio systems through primary user activity prediction. Radioengineering 21(4):1092–1100
8. Akyildiz IF, Lo BF, Balakrishnan R (2011) Cooperative spectrum sensing in cognitive radio networks: a survey. Phys Commun 4(1):40–62
9. Christopher Clement J, Bharath Reddy B, Emmanuel D (2016) An energy-efficient cooperative spectrum sensing strategy with robustness against noise uncertainty for cognitive radio networks. Arab J Sci Eng 41(9):3399–3405

# An IOT-Based Advanced Health Monitoring System

**Ghaatki Vaishnavi, Ramavath Kavitha, and E. Vijaya Babu**

**Abstract** IoT is fastest transforming technology in the healthcare market; in this, healthcare start-ups have become the thanking factor for all of us. Because of our busy schedules and our daily lives, monitoring the health of patient is quite a challenging task. Many people, especially those in their senior years, should be tested frequently. In these situations, we provide the ingenious target which is capable of automating the task with ease. The gadget utilizes a web server to provide a smart patient health tracking system which is capable of allowing the monitoring of patient health data such as heart beats per minute, oxygen levels in blood, human body temperature, room humidity, and temperature. From this model, a modernized IoT-based advanced health monitoring system using ESP8266 web server is proposed. An oxygen level measuring sensor also known as pulse oximeter sensor named MAX30100/102 monitors both the heart beats per minute and oxygen levels in blood of a patient. LM35 checks the human temperature. Similarly, DHT11 is a sensor that detects the room temperature and its respective humidity of a patient's room.

**Keywords** DHT11 · Enormous · Esp8266 · LM35 · MAX30100

## 1 Introduction

The world's population is rapidly growing. Cities that are accommodating a growing population suffer an extensive urban living pressure. Despite the fact that the overall needing resources in medication and provisions in bigger cities have been developing on a case-by-case basis, but still the supply is not being adequate and worth. The massive sheer on the city and the officials to look over the health care is substantially an awakening idea to technological advancements to find the appropriate solutions for the most challenging tasks.

---

G. Vaishnavi (✉) · R. Kavitha · E. V. Babu
Vardhaman College of Engineering, Hyderabad, India
e-mail: vaishnavibharadwaj9701@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_5

Modern technologies have increased in healthcare systems with respect to the provocative lives all over the world. We have seen an increase in attraction for wearing sensors in recent years, and thereby, those type of appliances have become more accessible and affordable to the users. Scientists looked advanced their ideas in using such miniatured devices that bring up an exotic outset in the applications of medication and health monitoring of such.

Doctors play an essential role in health check-ups in the traditional manner, which takes a long time for registration, appointment, and then check-up.

People nowadays procrastinate check-ups due to busy schedules which may absolutely result in delay. The proposed method can somehow overcome this challenge by reducing time. IoT is exploring industrial domains such as automation and control biomedical to improve health care. IoT is employed in personal health care facilities in addition to hospitals. This model helps us in detecting various factors like blood oxygen saturation levels, heartrate, body temperature, and humidity of the room.

It is an emerging technology that can help all of us get the next level of health care. It ensures that patients can carry or embed economical, dependable, and handy gadgets, allowing for errorless association. The detecting devices will lead to continuous capturing of signals, which will then be associated with key functional characteristics and relayed over a connectionless network. Now, the generated data will be saved, refined, and examined in conjunction with current medical information. From this, a better prognosis and advised early therapy with the allowed records of data can be made. This research allows the model to predict disease diagnosis even in the absence of doctor. Machines may also be coped to organize with required capsules and tablets based on a comprehensive review of medical databases, in addition to prediction.

One's physical well-being has always been a paramount factor and is a concern in which the advancements of technology have rapidly been increasing in. Just like the subtle castigation occurred in the recent years of the deadly virus COVID-19 has badly blasted various factors like economy. It is obvious that how the incident reflected awareness in terms of the importance of health care. In the outlook of certain state of affairs, it is absolutely the greatest thought to surveil the patient's prefect using remote health technology. So, the Internet of things is one of the best solutions to overcome the issue.

The applications in the vision of health care are one the most necessary in the field of Internet of Things (IoT) applications that have been made available to the globe. To say in a word, however, Internet of Things is being used very efficiently to connect modern medical resources and give individuals with appropriate and helpful healthcare services. Patients' health can be continuously monitored using modern sensors that can be worn or inserted in their bodies. The data hence collected can be made into a procedure of analyzing, aggregating, and mining in order to diagnose the disease in the early span of time. The processing algorithms assist clinicians in personalizing treatment and also makes sure to decrease the cost requirement in the health care with increase and beneficiary outcomes.

## 1.1 Literature Survey

Wearable monitoring sensors are given to the patients that measure ECG, blood oxygen levels, temperature, breath analysis, sweating, etc. [1]. Diseases can be determined with ease and involves huge machinery devices that may not be flexible to use. (not suitable for rural areas where there is no proper health care systems). From [1] requires a compulsory doctor's assistance. The disadvantages from [1] are that flexibility, affordance, and not suitable to manage everywhere. A lightweight access control system that may be used by healthcare IoT devices to provide delicate allocation, fast suspension, and transactional data confirmation.

The model [2] is the best typed with higher ordered features for cloud-based technology that we are aware of. To upload PHIs to HSP, sheer IoT devices solely use similar encoding. Furthermore, this study is the first to define and address the cloud reciprocity problem, [2] which is a possible security risk. The model comes up with a promising foundation for cloud-based PHI systems, according to the production ranking. WBAN—wireless body area network comes into the picture where it consists of wearable sensors for accession, end-to-end processing of data, transmission and reception, and association to the sensor layer. Whole recorded data is made out of a unit using a BT connection to the server. Also, only a couple of parameters like temperature and heartrate are measured in [3] Data is sent through the Internet here. The introduced model and [3] have a point common that both use Internet connectivity for data sending. For end-to-end secured connection start with huge assets restricted healthcare sensing devices, this method [4] presented a sheer secured protection protocol.

The computationally intensive processes are delegated to strongly establish in the vicinity of the healthcare sensing devices using a proxy-based technique. The protocol's energy consumption gives the best out of the daily basis. The whole description of the design, development, and testing of the depth sensor and feedback which is simple, light, and low-cost. As shown in prior work, [5] the model has shown to cope up with the limitations such as processing power, signal comprehension in a noise interference environment, and bulky gear. This model [5] was shown to detect the existence and setups faced in the navigation path, and let know the safest route for the safest navigation to the user during testing. The design is a combination of both CoAP and IEEE service/DIM which can be used on healthcare devices. In IoT situations, [6] this is employed. They also demonstrated a comparison of HTTP and 11,073 DIM performance with the help of CoAP.

They also used CoAP to assess performance. When it comes to the profusion of packets in a single HTTP request, using JSON, [6] you may track transactions, packet loss rate, and syntax, as well as XML. Finally, they came to the conclusion that CoAP is capable of in comparison with HTTP, only a few packets are transmitted. In terms of quality, exploitation of resources, they claimed that XML is inferior than HTML compared to JSON.

In the cloud, the health parameter data of individuals is properly stored with all kinds of security and protection, as well as the reduction of hospital stays for routine examinations. It's critical that health and sickness can be tracked. Any doctor, at any distance, can diagnose you.

An Internet of Things (IoT) device is used in this paper [7]. It is created in such a way that the system which monitors the health is purely based on web. The system is in place. Temp of body, humidity of the room, and rate at which heart beats measured with sensory technique is used for temp measurements. The sensor data is subsequently transferred to a medical server through the Internet. Communication over the airwaves after then, the data is received in an electronic format [7]. Smart phone, personal ones, is provided with lot platform. With the help of when the doctor receives the results, he or she diagnoses the condition. The study [8] carefully summarizes and expands on existing evidence and practitioner literature on big data platform implementation and mHealth outcomes. IoT wearables. In the context [8] of the healthcare industry, which is dominated by women, the focus is on patient-centered solutions.

Implementation of a big data platform with the establishment of a big data platform, it is feasible to investigate diverse, autonomous, ongoing, and changing systems information (HACE). The amount of data gathered by the sensors will be enormous. HACE will be followed by the data. Big data, on the other hand, can be beneficial, handle a lot of data, a lot of variation, and a lot of speed. It is possible to attain veracity. Implementation [8] of a big data platform is a measure of how much big data is being used by healthcare providers, for storing and analyzing patient information. It is difficult to foresee the chronic disorder in initial stages. The advantages of an IoT-enabled remote health monitoring system over traditional health monitoring are significant system. It is critical to collect accurate raw data in a timely manner. However, [9] it is more important to investigate and the resources. Raw data in order to extract additional valuable data, such as connections between things and services to allow for a web of things or a service-oriented Internet. Darshan and Anandakumar [9] IoT in the healthcare system, as well as a review of previous work carried done on this study topic for which a proposal has been made. Notification from e-mail which give alerts signals, checking through the website in Internet fir monitoring of remote data as well as the name and dosage of any prescription medicines.

Display of doses It has the potential to reduce human error. The majority of the health state of the participants is an important component of this system. The patient could also be observed from the comfort of his or her own home. During a semi-major illness, essential action could be performed. The possibility of human mistake during data collection could be high. As sensors are utilized to collect health data, this [10] will be effectively decreased. Because email alert notification and website visits for remote data monitoring through the Internet are dependent on the correct capacity of the Internet, this system requires a sufficient amount of bandwidth. The suggested [11] system's goal is to determine acceptability, validity, and usability. As a result, the technology has been put to the test with real patients in a medical setting at Khulna University. Khulna Medical Center is located in the Bangladeshi city of Khulna. The outcome of the experiment is given here, along with a photograph taken during the experiment. The acceptance, validity, and usability tests are all assessed by the same person. A few survey questions additionally, data transit and response time.

The system [11] is scrutinized. With the advancement of technology, it is to be expected that Internet access would be available everywhere. Furthermore, low-cost sensors are simple to integrate with any system. To provide global information service in addition, these sensors have less power consumption and ability to execute work as well as larger machine, machineries for measuring health. We've attempted to create one. A technology that can be used to remotely monitor a patient's condition taking the necessary efforts to improve one's health. The purpose of this system will collect data from patients in an automated manner.

The Internet of Things (IoT) technology has recently advanced to the point where all objects are now connected. We might claim that the Internet of Things (IoT) has been identified as the next technological revolution. To name a few, [12] Internet of Things applications include smart parking, smart homes, smart cities, and smart environments in industrial settings, as well as irrigation and health monitoring. The healthcare industry is one of these applications, since patient health metrics are monitored. The efficiency of medical equipment can be monitored and improve by Internet of Things by monitoring of patient's health in real time where specific sensors collect data from the patient. Because this technology digitally senses and transmits patient health data, it decreases manual involvement and, as a result, human errors. Patient parameters are transmitted through medical equipment via a gateway, where they are stored and evaluated in the Internet of Things. Kubal and Chandrashekarappa [12] the monitoring of all patients from multiple locations is one of the major obstacles in the introduction of Internet of Things for healthcare applications. Thus, in the medical industry, the Internet of Things provides a solution for effective patient monitoring at a lower cost, as well as a reduction in the trade-off between patient outcome and disease management. The biomedical parameters are saved on a website [13] that can be utilized as a medical history for the patient in the future.

Several patients can have their gadgets registered. To the server, and can use accounts to retrieve their settings logging in (each account has a distinct threshold value set by the user) after a thorough examination by a doctor. When the value of the currency is when a physiological parameter surpasses a predetermined value, the doctor or career is notified by SMS and e-mail. This a reliable method that can be used to keep track of a person's health. Patients who are elderly or unwell who are unable to visit the hospital on a regular basis. In emergency situations, they demand constant assistance. This system's goal is to make it portable, so it [13] won't get in the way. The ECG electrodes were implanted in the appropriate location to monitor the patient's ECG, and the [14] obtained signals were transferred to IBM Bluemix cloud. The genuine article on the Bluemix dashboard, time ECG graphs were produced for a set amount of time the value of the ECG signal at any given time. Gauge was also used to display the information.

## 2 IOT in Healthcare System

IoT has been one of the rapidly growing branches in the sector of Data Science and Information Technology. It has become the progress for wireless communications recently, and is extensible on Internet. It has capability to connect to some extent of sensors to the Internet like infrared, identification (radio frequency), etc. Among, IoT has a pivotal part in Medicare. The world has now become to where a lot of people lose their lives due to many various diseases. In fact, the deadliest virus called COVID-19 has already been spread and is being spread.

According to some reports, people dying with chronic diseases are on an average of 60 grown while people are putting efforts in maintaining a healthy life, IoT helps as the common point in the field of health care systems. IoT links to the physical world. Here, physical world means household appliances like, thermostats, human body, construction, etc. On initiating some of the IoT technologies in this sector will literally progress the quality of people's lives and their personal knowledge. Some of the applications are as follows:

**Keep track on health**: There are some various wearable devices that are capable of detecting basic activities like stress management, BPM and health-based progress like apple that provide smart watches. This may lower the source of supply but is very different with respect to other devices that are present in hospitals.

**Health information uphold**: IoT helps in referring medicine consumption from time to time. Devices like ECG, pulse oximeter and sphygmomanometer can improve the measurement and monitoring resulting in accurate outcomes.

**Support and Service**: IoT is capable of connecting network systems and vehicles like cars. For suppose, if someone meets with an accident, the location and direction of the tragedy can be recorded and sent to the respective support systems like traffic sector and emergency where in one has a possibility to get well.

**Big Data collection**: IoT is capable of generating big data health. The enhancement is done with respect to the progress and growth of health care systems in IoT. It consists of 5 and 3.3 V supplies. The 3.3 V is used to supply for heartrate sensor (MAX30100) and ESP8266 Wi-Fi module. The power supply can be given from gadgets like, laptop, power bank, etc. (Fig. 1).

It consists of 5 and 3.3 V supplies. The 3.3 V is used to supply for heartrate sensor (MAX30100) and ESP8266 Wi-Fi module. The power supply can be given from gadgets like, laptop, power bank, etc.

The block diagram shown above has several temperature sensors; the temperature once blocks, they can be categorized as Arduino UNO board, regulated power supply, IoT-based Wi-Fi module, buzzer, smart phone, and sensors.

1. Arduino UNO: Arduino plays a vital role in the description of the design. Here, the Arduino Uno-based microcontroller board is typically used for processing and control unit.

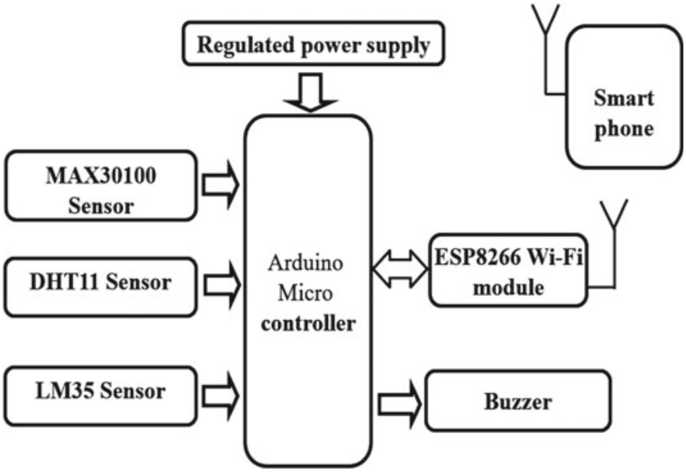


**Fig. 1** Block diagram

2. Regulated Power supply: This block is mainly for supplying power to the Arduino-based microcontroller board and other modules as well. This regulated power supply consists of 5 and 3.3 V supplies. The 3.3 V is used to supply for heartrate sensor (MAX30100) and ESP8266 Wi-Fi module. The power supply can be given from gadgets like, laptop, power bank, etc.
3. ESP8266 Wi-Fi module: The ESP8266 is an IoT-based WI-FI module which works on TCP/IP integrated protocol stack. The role of ESP8266 here in the design is to work as a Wi-Fi hotspot with a particular SSID and password in order to check the respective parameters from the default browser obtained from the code.
4. Sensors: The sensors in the design are used to detect the body parameters like human body temperature, room humidity and temperature, spo2 and heartrate of a human body. They are connected with the Arduino board. The sensors used in the design are MAX30100, LM35, and DHT11. Max30100 is capable of detecting heartrate and spo2 levels in a human body. The LM35 is capable of detecting the human body temperature. The DHT11 is capable of detecting the room humidity and temperature of the patient staying. The sensors act as the input modules in the design, while the heartrate sensor of 3.3 V operates interface using I2C serial communication. In case of body temperature sensor, the temperature once recorded is converted into the equivalent voltage. For this, the LM35 module operates at 5 V DC.
5. Buzzer: The buzzer in the design beeps at the proportionate of the heartrate of the human body.
6. Smart Phone: The smart phone plays an important role as to record the values of the parameters checked from the sensors with the default browser obtained from the webserver.

An ESP8266 is an IoT-based model of Wi-Fi. The ESP8266 is an integrated chip which consists TCP/IP. This permits access to any Wi-Fi coverage through any microcontroller. The ESP8266 is basically a dual mode operator. The two modes it can carry is unloading Wi-Fi functions and application hosting. The advantage with ESP8266 is that it comes pre-programmed internally with some of its AT commands which will obviously make the work easier to directly connect it with the Arduino and can utilize the ability of the Wi-Fi.

ESP8266 module is a bit expensive with respect to its growth and community usage which is often the rarest disadvantage of it. The module itself calls it as integrated due to its limitless processing and storage capacity through the GPIOs it has. Its degree is permitted to short outraging circuitry along with the front end which is often designed to occupy the smallest area on the PCB. The ESP8266's RF adjustment permits it to work under all circumstances, and hence, no extra part of the RF is required (Fig. 2).

The schematic diagram that is referred above helps in explaining the interfacing of every component with the sensor that measures heartrate and the microcontroller. Firstly, the regulated power supply of + 5 V has to be given to the Arduino port by any means of USB connector.

The LM35 sensor has to make a connection with the A0 pin in the Arduino board. Similarly, the DHT11 sensor that is used for room temperature and humidity has to make a similar job to the A2 pin of the Arduino. Max30100 sensor has to be connected to the A4 and A5 i/o pins that is used to measure both heartrate and spo2 levels. Finally, the ESP8266 Wi-Fi module has to be connected to the A0 pin. A

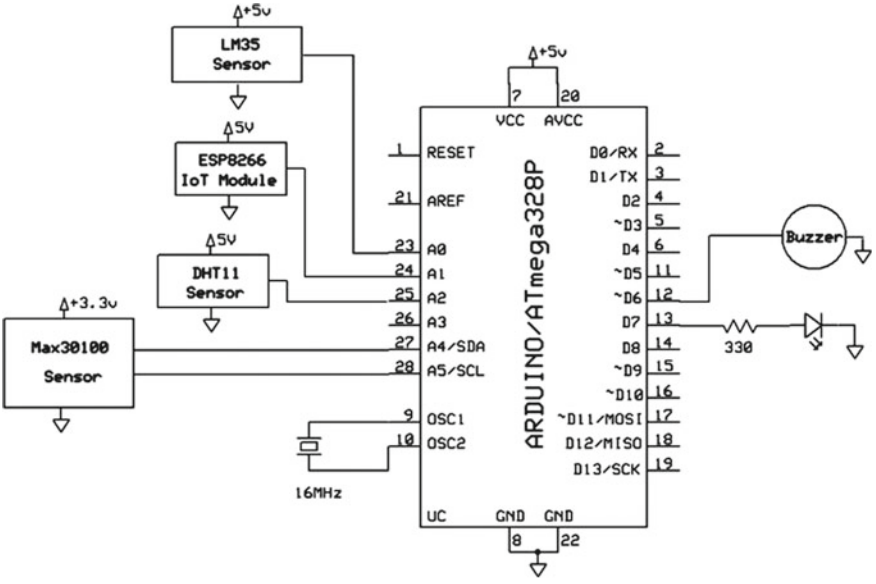


**Fig. 2** ESP8266 module

buzzer is also placed at D6 pin of the Arduino board to proportionate the sensors. This work can be implemented using the following set of software:

1. Express PCB—for designing circuit PIC
2. C compiler—for compilation part
3. Proteus 7 (Embedded C)—for simulation part in order to obtain the IP address apart from the default browser available for the ESP8266, we need to cross check and compile the code in the Arduino software for that particular IP address.

## 3 Results Obtained

So, in this model, while obtaining results from the web server, we have got numerous variations of parameters with respect to different health conditions of every person disease distributions and patterns (Table 1).

The values have shown up to be accurate while measuring all of the factors like heartbeat rate, oxygen spo2, body temperature, room humidity, and room temperature. We have also cross checked with different rooms in case of room humidity and temperature (Fig. 3).

**Table 1** Data results

| Parameters | Values |
|---|---|
| Heartrate | 92 BPM |
| Oxygen Spo2 | 95% |
| Body temperature | 80° |
| Room humidity | 20% |
| Room temperature | 23° |

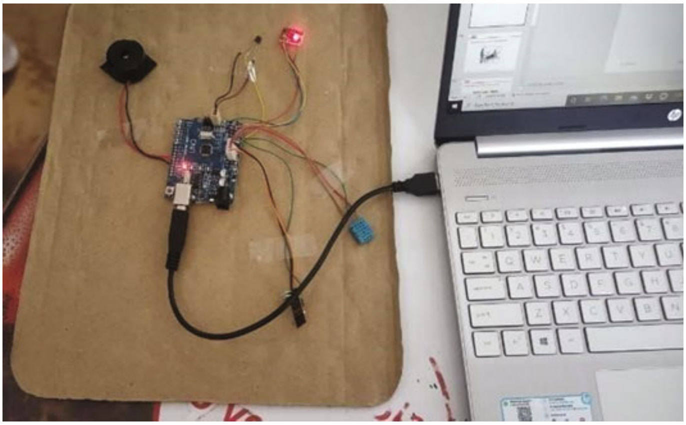


**Fig. 3** Project setup

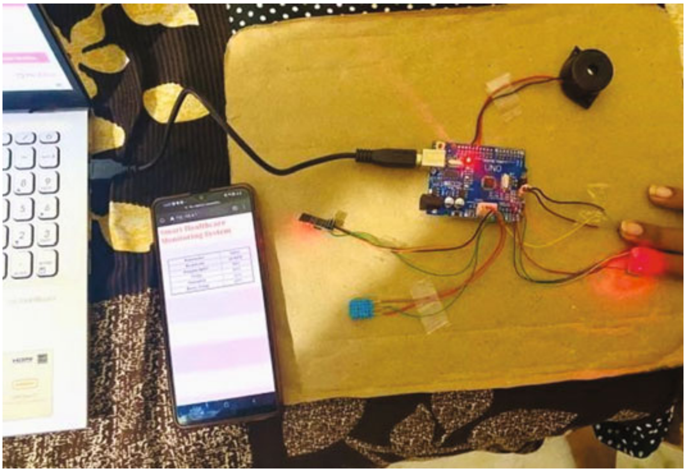


**Fig. 4** Project setup along with parameters data

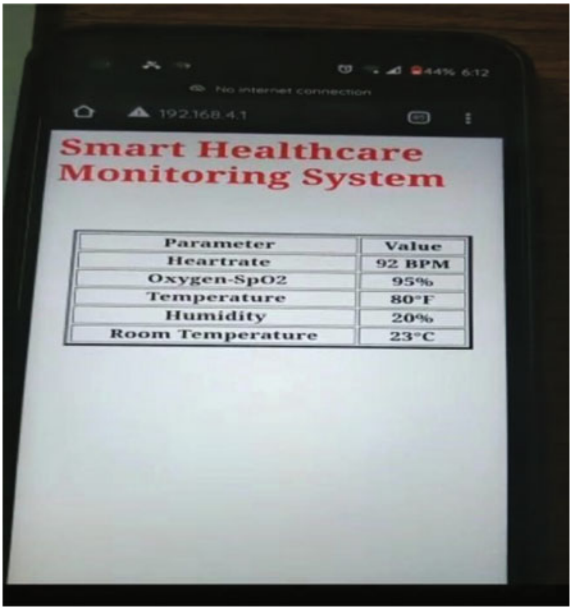
Hence, it’s been concluded that this method is often called an efficient way to diagnose the disease quickly. The obtained results from the web server can be shared over a network by an IP address to someone who is close to the patient, by which the chance of intimating happens (Fig. 4).

Throughout the system design, we noticed various variations that helped us tracking the right results. This technique is very much helpful in the conditions like outbreak of COVID-19 where there is no possibility of contact from one person to other. Thus, designing this system has many advantages like, non-contact hygiene, affordance, ease of use, and one’s self knowledge to check his/her own personal check of the body. This model gives the overall flexibility to the user and one can recommend (Fig. 5).

While discussing the advantages and disadvantages of the design, features like, accurate heartbeat rather than the shown one, apt and cost-free design, quicker response, less power consumption and ease of design and installation come under the advantages. Drawbacks like limited distance and sensitive interfacing with respect to heart beat sensor to the microcontroller are post valid to the future scope of the design.

## 4 Conclusions and Future Scope

The vital features of the components used are designed and developed in an efficient manner. The design made reasons for each and every component used. Also, the usage of higher order functionalities with the ICs has made the design successful. So, finally here, we reviewed various parameters check using branch of data science ‘IoT’ along with the help of ESP8266 web server. The body temperature, blood



Smart Healthcare Monitoring System

| Parameter | Value |
|---|---|
| Heartrate | 92 BPM |
| Oxygen-SpO2 | 95% |
| Temperature | 80°F |
| Humidity | 20% |
| Room Temperature | 23°C |

**Fig. 5** Data results obtained on phone

pressure, spo2, room humidity, and temperature are prime parameters to diagnose the disease, and hence, we can conclude that IoT is being rapidly increased in various applications out of which healthcare are one of them. To let technology, move a step forward with the branch of data science indeed makes it a way more efficient and ease for use. We highlighted comparative analysis and visualization of data using the web server ESP8266 for more accuracy and flexibility. For more future analysis, affordance for every user increases with the increase in usage. This project can also have a scope of extension by using RS485 that may make use of wired supply but that can somehow increase the distance. Many wireless technologies like Zigbee can be used that improves the ease of using wired connections.

Meanwhile, the doctor's role is crucial as to pay more of an effort in a long run to encourage patients and help them to make use of RPM. Needless to say, the drawback of this technology is that the substantial accuracy of the devices used. More the imprecision, more the effect of RPM exists.

## References

1. Khan SF (2017) On IoT for healthcare systems in the essence of RFID. In: International conference on industrial technology and management, IEEE
2. Yeh LY, Chiang PY, Tsai YL, Huang JL (2015) Cloud-based fine-grained health information access control framework for lightweight IoT devices with dynamic auditing and attribute revocation. IEEE Trans. Cloud Comput. PP(99):1–13

3. Page A, Soyata T, Gaurav Hassanalieragh M (2015) IoT in healthcare using cloud-based monitoring. Opportunities and Challenges
4. Porambage P, Braeken A, Gurtov A, Ylianttila M, Spinsante S (2015) Secure end to end communication for constrained devices in IoT-enabled ambient assisted living systems. In: IEEE 2nd world forum on internet of things, Milan, pp 711–714
5. Yelamarthi K, Laubhan K (2015) Navigation assistance and space perception “for visually impaired with depth sensor and haptic feedback. Int J Eng Res Innov 7(1):56–62
6. Joyia GJ, Liaqat RM, Farooq A, Rehman V (2017) Internet of medical things (IOMT): applications, benefits and future challenges in healthcare domain. J Commun
7. Banka S, Madan I, Saranya SS (2018) Smart healthcare monitoring using IoT. Int J Appl Eng Res 13(15):11984–11989. ISSN 0973-4562
8. Perumal K, Manohar M (2017) A survey on internet of things: case studies, applications, and future directions. In: Internet of things: novel advances and envisioned applications, Springer International Publishing, pp 281–297
9. Darshan KR, Anandakumar KR (2015) A comprehensive review on usage of internet of things (IoT) in healthcare system. In: Proceedings international conference on emerging research in electronics, computer science and technology
10. Saha J (2018) Advanced IOT based combined remote health monitoring, home automation and alarm system
11. Ghosh AM, Halder D, Alamgir Hossain SK (2016) Remote health monitoring system through IoT. In: 5th International conference on informatics, electronics and vision (ICIEV)
12. Kubal A, Chandrashekarappa K (2017) Design of EHealth monitoring of patient using internet of things
13. Hagargund AG, Srivastav A (2018) Smart and automatic health monitoring of patient using wireless sensor network
14. Singh P, Jasuja A (2017) IoT based lowcost distant patient ECG monitoring system. In: International conference on computing, communication and automation (ICCCA)
15. Alnosayan N, Lee E, Alluhaidan A, Chatterjee S, HoustonFeenstra L, Kagoda M, Dysinger W (2014) MyHeart: an intelligent mHealth home monitoring system supporting heart failure self-care. In: eHealth networking, applications and services (Healthcom), 2014 IEEE 16th International Conference on, IEEE, pp 311–316

# Fault Tolerance in Vehicular Cloud Networks: Link Failure and Malicious Node Detection and Reconfiguration

**Sandeep N. Kugali, Bibhu Kalyan Mishra, Sneha Kadadevar, Piyush Kumar Pareek, and Malaya Malla**

**Abstract** Vehicular cloud networks (VCNs) are becoming more popular these days due to their applicability in multifarious applications across the globe. The VCNs are alternative application of the cloud computing domain that is an arrangement of the cloud as well as Vehicular Ad Hoc Networks (VANET) technologies. Primarily, the VCNs are made of diverse clouds namely the infrastructure cloud, conventional information technology (IT) cloud, and vehicular cloud. Failure of link and malicious node detection are becoming a major issue nowadays due to the increasing traffic across the globe. Some of the major problems with VANET are routing, security, and privacy that demands more attention and pragmatic solutions in order to minimize the chances of accidental situations on the road. In this paper, a novel model is proposed in order to resolve the existing issues related to the fault tolerance in the VCNs, i.e., link failure and malicious node detection and reconfiguration. This test was performed on the CloudSim simulator on a personal computer having 64-bit operating system and 8 GB (RAM). The results are observed with a high degree of precision and compared with Heterogeneous Earliest Finish Time (HEFT) algorithm in order to validate the

---

S. N. Kugali
Department of Information Science and Engineering, Basaveshwar Engineering College (Autonomous), Bagalkot, Karnataka, India
e-mail: snkis@becbgk.edu

B. K. Mishra (✉)
FOS, Controller of Examination, Sri Sri University, Cuttack, Odisha, India
e-mail: bibhu.m@srisriuniversity.edu.in

S. Kadadevar
Department of Computer Science, Basaveshwar Arts Science and Commerce Independent PU College, Bagalkot, Karnataka, India
e-mail: snkis@becbgk.edu

P. K. Pareek
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, India
e-mail: piyush.kumar@nmit.ac.in

M. Malla
FCMS, Sri Sri University, Cuttack, Odisha, India
e-mail: mm2@srisriuniversity.edu.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_6

proposed model. Although, there has been carried out extensive examination in this sector during the last decade, but there is a pragmatic scope of more research in order to explore the full potential of this domain.

**Keywords** Link failure · IoT · Node · Traffic · VANET · VCNs · Vehicle · Dynamic architecture

## 1 Introduction

The Vehicular Ad Hoc Networks (VANETs) are considered one of the biggest real-life applications of multifarious popular ad hoc networks wherein the nodes may be represented with the help of fast traveling vehicles [1]. Nowadays, smart cities are becoming one of the best methods of viable urban expansion that is made of the sensor networks as well as Internet of things (IoT) and many more components [2]. Secure transportation is becoming tougher in large cities where traffic is in huge amount and demands more attention toward new models and systems in order to effectively maintain the traffic and make roads safer during traveling [3]. Intelligent transportation systems (ITS) are in high demand nowadays across the globe form the safety perspective of automobiles on road during traveling. The VANET is primary element of the ITS that may increase the overall efficiency of exchanging information over the road's traffic [4, 5].

VANET is different class of network where each automobile may exchange on-road information namely vehicle accidents or the jam by vehicles on road onward [6]. Communication can take place in two separate ways in the VAN ET, one is the vehicle-to-vehicle communication and another way is vehicle-to-infrastructure communication, where each automobile separately establishes a communication link with each other through the existing infrastructure over a road or highway accordingly [7]. Figure 1 shows the diagram for the VANET infrastructure. Figure 2 shows how VANET makes safer roads and assure fewer or no accidents. Figure 3 shows how VANET provides more efficient driving and letting the d river know about the traffic. Road safety is becoming one of the major issues in various areas nowadays and demands effective solutions in the modem world. The VANETs demand higher mobility as well as position awareness of automobiles which participate in communication processes [8]. All automobiles must have accurate location coordinates of all automobiles within network in order to handle any emergency condition [9]. There would be huge no. of heterogeneous intelligent automobiles. Effective management of heterogeneous automobiles as well as sporadic relations is critical challenge now a days in the arena of VANETs [10]. There are gigantic demand of several sensors that must be integrated within the automobiles, and cloud would acquire as well as process such acquired data prior to allocation with certain portion s of networks [11].

The VANETs usually utilize a mutual medium in order to establish a communication link, and therefore, media access control (M AC) scheme drawing has been

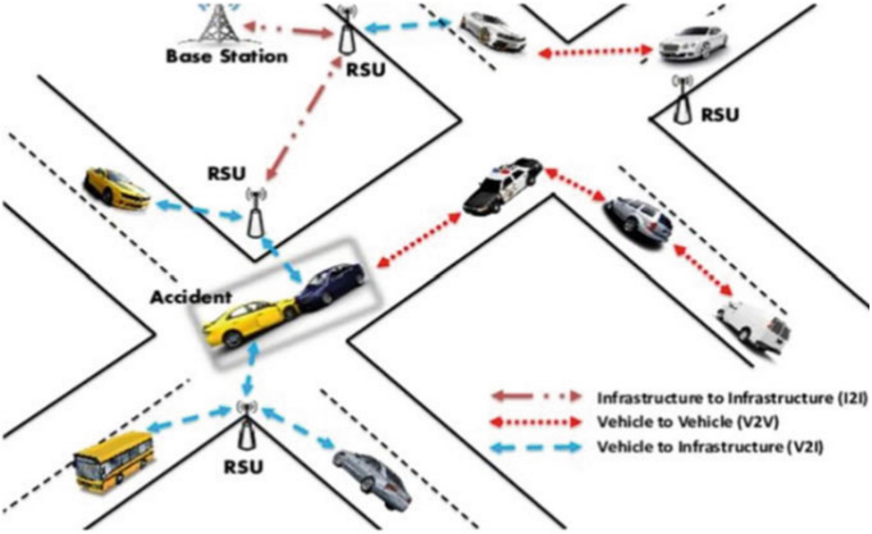


**Fig. 1** The diagram for the VANET infrastructure [12]

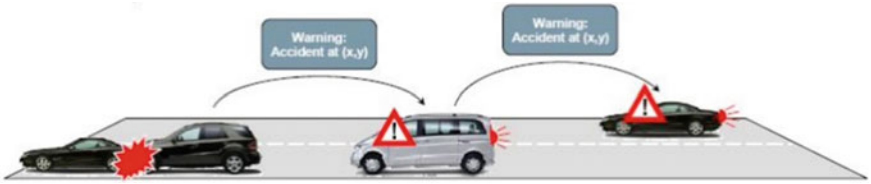


**Fig. 2** The VANET promises safer roads, assure less or no accident

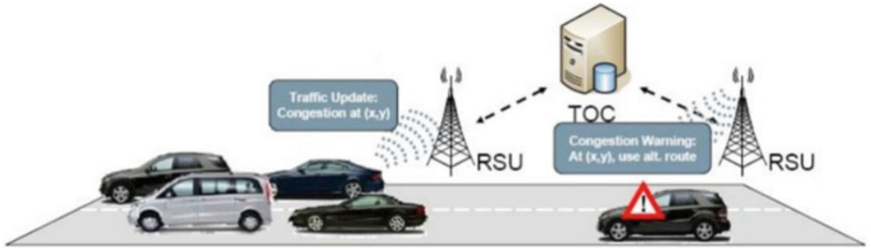


**Fig. 3** How VANET provides more efficient driving and letting the driver know about the traffic

considered major problem nowadays [13]. There are multifarious technologies investigated during last decade namely Time Division Multiple Access (TDMA), as well as Carrier Sense Multiple Access (CSMA) and many more; IEEE 802.11 accepted CSMA rooted MAC for many applications of the VANET [14]. There are encountered multifarious essential task in the sector of VANETs in order to maintain Quality of Service (QoS) [9]. There is still an enormous requirement of such kinds of the networks that reduces the delay in more pragmatic way as well as highest connectivity time offers the QoS to the users [15]. For proper transmission and reception of the data packets via various nodes, a more efficient as well as pragmatic algorithm is highly demanded in modern world.

In the VANETs domain, any pragmatic routing algorithm is an optimal routing method that has very less delay, mini mal capacity must be required for system as well as lower computational complexity. There is a conducted multifarious investigation by the investigators in the field of VANETs for implementations of novel methods in numerous networks topologies for satisfying each property in a pragmatic way [16, 17]. Figure 4 shows the multiple applications of VANET in fun and entertainment such as text messaging and MP3 download. Figure 5 shows the diagram of the smart vehicle. There are installed multifarious components such as positioning systems and many more as depicted.

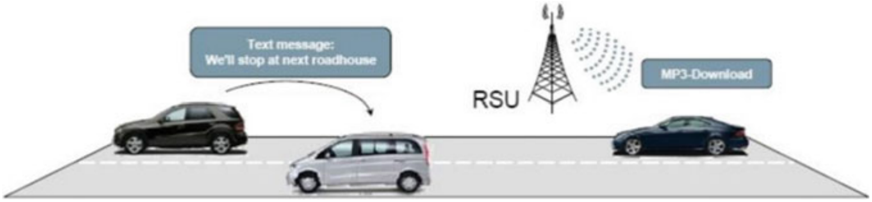


**Fig. 4** The multiple applications of VANET in fun and entertainment such as text messaging and MP3 download

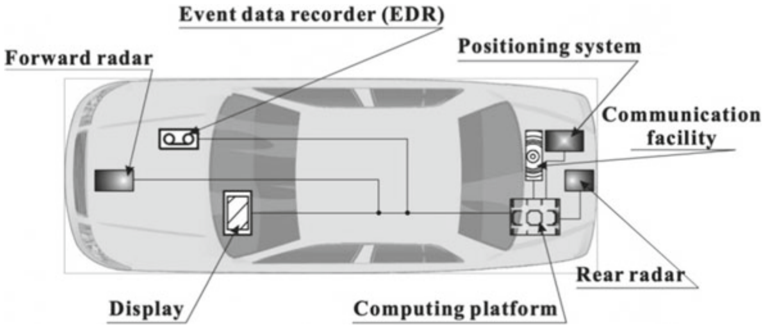


**Fig. 5** The diagram of the smart vehicle. There are installed multifarious components such as positioning systems and many more as depicted

## 2 Literature Review

Zeadally et al. carried out an investigation on Vehicular Ad Hoc Networks (VANETs) in their research paper. Rapid innovations in infrastructure, applications, and networking technology have made it possible to build and introduce a wide range of diverse kinds of networks in variety of settings. VANETS are one specific framework that has attracted more attention d u ring last decade. The VANET is becoming one of the most dynamic field for research as well as deployment and standardization as this comprising huge potential in order to enhance the vehicles as well as road protection, improves the overall efficiency of the traffic, and provides much ease to the drivers and all the passengers. Enormous research has been performed in this field in recent years in order to explore pragmatic ways of existing issues encountered by multifarious researcher during last decade. In this paper, authors provide detailed overview on the VANET and its applications in various fields as well as offers advantages over other networks [18].

Yang et al. conducted a research in the field of vehicular network in their research paper. In this paper, authors discussed about the impact of broadcasting in vehicular networks as it is becoming one of the pragmatic fields due to huge demand by researchers during the last decade. It is becoming very tedious task to deliver a message to multifarious vehicles instantly due to the inefficient network bandwidth and traffic congestion over the wireless network. Such problems are becoming very common nowadays a demand more attention to develop novel models and infrastructure in order to efficiently handle such critical situations in most efficient manner. In this paper, authors highlight the current issues of link failure during the emergency and critical situation when communication is necessary in emergency situations [19].

Chen et al. explored a review paper on the topic of VANET in order to highlight the current routing problems. This paper provides a detailed overview on the VANET current problems related to the link failure during the communication and how it can be resolved in pragmatic manner in the future in order to improve communication among several automobiles over the road to avoid accidental situations immediately. This survey paper is a detailed examination of existing issues encountered in arena of VANETs by multifarious researchers. Moreover, authors describe the main ideas of the current technologies that could be beneficial for the researchers in order to develop the novel model and methods in less lime without comprising huge and costly infrastructures [20].

## 3 Methodology

### 3.1 Design

Vehicular Ad Hoc Networks (VANETs) are considering hot exploration domain during last some years. There are multifarious separate characteristics namely highest dynamic topology as well as movement. The VANETs are attracting huge attentions toward more research in order to determine and resolve existing threats in a very pragmatic manner. It is a special form of MA 1 T and it provides vehicle-to-vehicle communications and vehicle-to-infrastructure communications. Nowadays, VANETs are utilizing armed automobiles as the network nodes. The controlling as well as effective administration of network links surrounded by automobiles as well as setup is becoming key encounter. Some of intermittent links because of highest movement of automobiles or higher packet damage within vehicular links should be escaped. There are multifarious threats and risk challenges that are related to secrecy of multiple users data as well as their positions in certain areas. There is a need to develop and implement some kinds of the infrastructure that must permit multiple users for deciding which kind of the data or an information must be shared and what must be reserved. Figure 6 shows the Vehicular Ad Hoc Networks (VANETs) architecture. Figure 7 shows the design of the proposed model. Figure 8 illustrates the message propagates to the destination using a number of intermediate links. Figure 9 shows if vehicle mobility causes links to break, message rerouted using a different path.

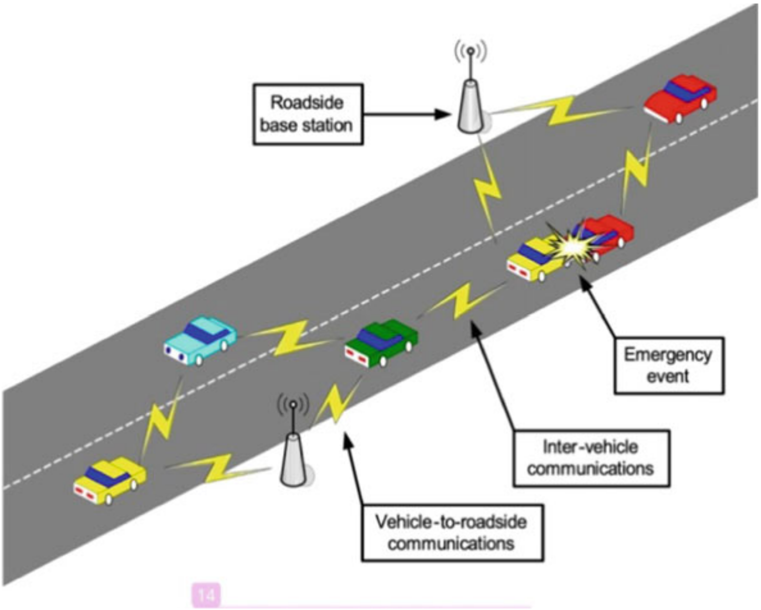


**Fig. 6** The vehicular ad hoc networks (VANETs) architecture

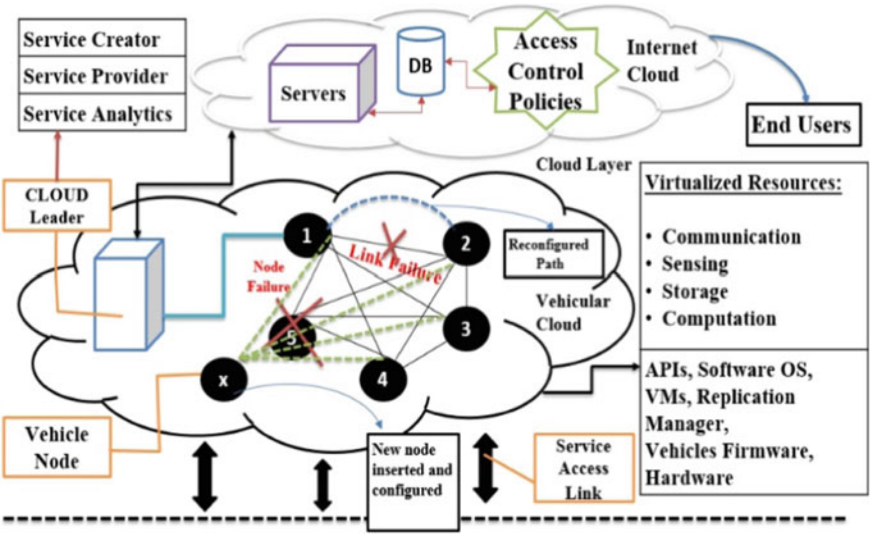


**Fig. 7** The design of the proposed model

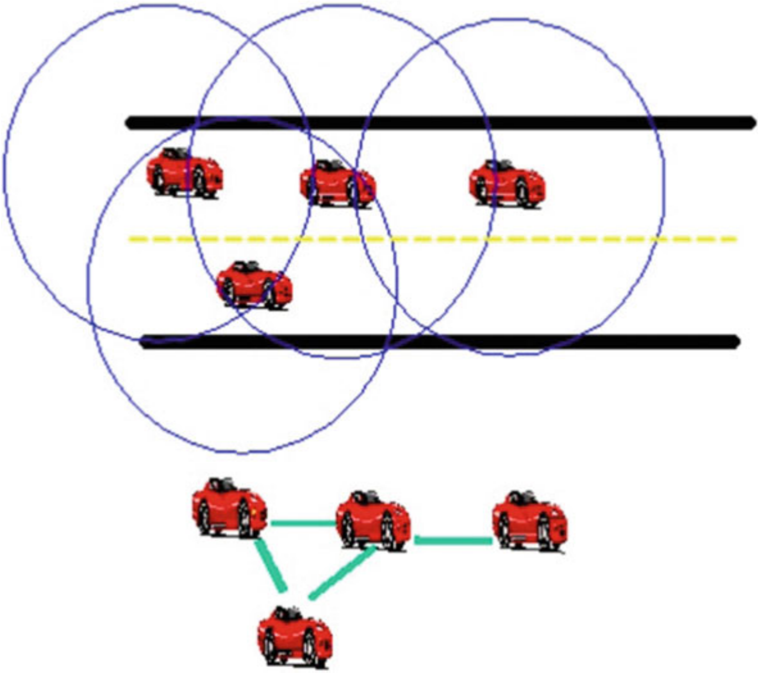


**Fig. 8** Message propagates to target utilizing a number of middle links

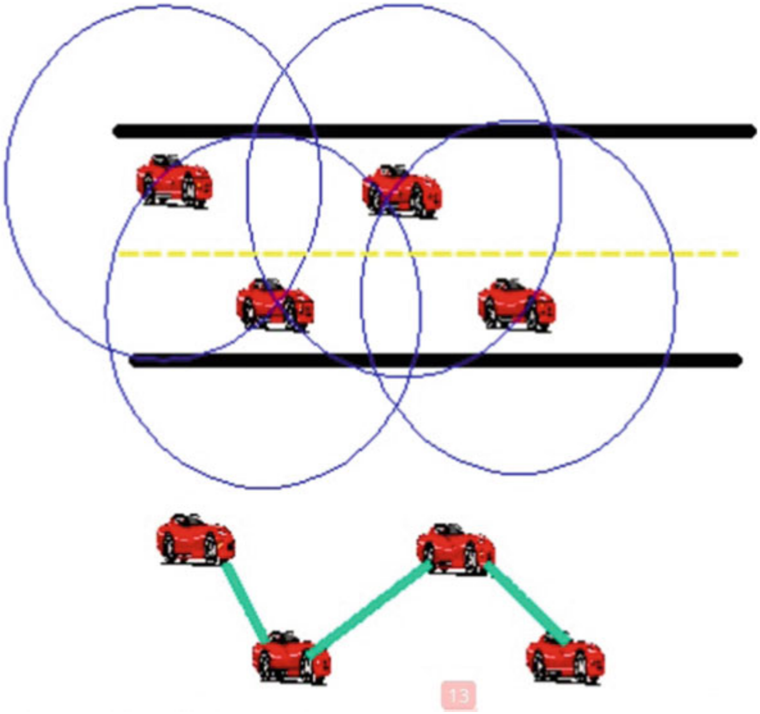


**Fig. 9** If vehicle movement cause links to disruption, message rerouted using a different path

## *3.2 Instrument*

This experiment was conducted on CloudSim simulator installed on a personal computer having a system configuration of 64-bit operating system, 8 GB RAM, and Windows 10. The simulation procedure was carried out with a high degree of precision to validate the measured results in an efficient manner. CloudSim simulator is basically a framework that is utilized for certain works such as modeling or simulation process of the cloud infrastructure. The CloudSim simulator is open source software utilized globally for the testing of cloud infrastructure or multifarious services. This simulator was written in the Java programming language. This software allows developers as well as scientists to test several cloud applications according to the requirement. The CloudSim is one of the most pragmatic and easy to u se simulation toolkit nowadays which supports modeling as well as simulation of main functionality of the cloud platform, namely tasks line, event process, cloud entities creations, message among diverse entities, and several others. This software allows to the users for testing of application services in repeated as well as in a controlled environment.

**Table 1** The selected parameter values for the simulation procedure of the experiment

| S. No. | Parameter | Value |
|---|---|---|
| 1 | No. of the vehicle | 4–25 |
| 2 | No. of the tasks | 25 |
| 3 | Startup time | 3 (s) |
| 4 | Workload | 15–65 |
| 5 | Communication capacity | 6–12 (Mbps) |
| 6 | Task degree | 0–2 |
| 7 | Computing capability | 4–9 (MIPS) |
| 8 | Velocity | 8–18 (ms) |
| 9 | Traveling distance | 0–150 (km) |
| 10 | Service failing probability | 0–1.0 |
| 11 | Data of communication | 1–35 (Mb) |
| 12 | Radio boundary distance | 3000–5000 (m) |
| 13 | λ1 | $0.90 \times 10^{-5}$ to $1.10 \times 10^{-5}$ |
| 14 | λ2 | $0.00–10^{-4}$ |
| 15 | λ3 | 0.06 |

### 3.3 Data Collection

The selected parameter values for the simulation procedure for the proposed model are illustrated in Table 1. The total number of the vehicles were considered in between 4 and 25, and no. of the tasks were taken 25 for the testing purpose. The startup time taken 3 s initially and communication capacity were 6–12 Mbps. The testing procedure was performed with high security in order to eliminate the chances of the error and to validate propped model accurately.

### 3.4 Algorithms

There has been done extensive development in the field of wireless communications constantly with time. The VANETs are becoming one of the demanding and pragmatic technologies in the arena of wireless communication in comparison with other existing technologies. The VANETs are playing a key role in modern communication systems because of the faster handovers, availability of the network, secrecy, robustness, safety, and many more features for advanced applications. The VANET technology has grown but certain issues still demand more attention in order to make networks more vigorous. In this paper, novel algorithm s are investigated for the route discovery (in case of link failure), route maintenance, node failure as well as for node replacement in an efficient manner in comparison with existing HEFT algorithm. The HEFT algorithm has various drawbacks with modem communication systems such as inefficient distribution of the task and many more. The primary notion behind existing HEFT set of miles is that one rank is to be evaluated for each task within Directed Acyclic Graph (DAG). The existing HEFT algorithm works in two diverse segments; initially, the rank value must be evaluated, and in later segment, the processor task is completed. In order to resolve the aforementioned drawbacks, new algorithms are proposed in this paper. The fault tolerance within the VCNs, i.e., failure of the link and malicious node detection as well as reconfiguration is becoming a critical issue nowadays.

- ***Route Discovery Algorithm***:

  Rreq: Route Request, Rrep: Route Reply, : all nodes, Nd: destination node, s: source node

  ```
  lf pkts are transported for source to target
  {
  Broadcasting Rreq. to N
  if ($Ns$ = Nd)
      {
          Send Rrep to Ns
      }
  Else
      {
          Broadcast Rreq to neighbor nodes
      }
  }
  ```

- ***Route Discovery Algorithm (considering Link Failure)***:

  1. If pkts are available to forward
  2. Until destination node is not found
  3. If no link failure present
  4. Add current node inside the route
  5. Else if link fails
  6. Find the neighbor node and send Rreq
  7. If destination node found send Rrep

- ***Route Maintenance***:

  Rerr: Route error

  1. If there are nodes failure, then inform the Ns through sending Rerr
  2. Create a fresh parallel track from Ns to the Nd
  3. Handover these pkts which are in buffering at the time of connection failure

- ***Node Failure Detection Algorithm***:

  d1: time before taking decision that node is failing d2: time after taking decision that node is failed

```
1. Transmit a small message say notice
2. Waiting for d1
3. If no reply from neighbor nodes then after
4.     Neighbor node is dying
5. End if
6. Pause for d2
7. Drive to 1
```

- ***Node replacement Algorithm***:

```
If dying nodes are determined then
    For every N belonging to directed links do
        Transmit data message of movement
    End for
    Find any extra node, end if
    If chosen auxiliary then
        Display to safeguard the functions dying node
        If its directed neighboring the list is blank then
            Continue some time
            Arrival to its location
        Endif
    Endif
```

Vehicular Ad Hoc Network (VANET) is becoming most demanding technology nowadays across the globe as it resolves the existing issues of link failure and malicious node detection and reconfiguration in an efficient manner. In this paper, a novel model is presented for fault tolerance in vehicular cloud networks in order to reduce the average execution time as required in present systems due to technological advancements. This testing procedure was executed on the CloudSim simulator with a personal computer that has a system configuration, 64-bit operating system, 8 GB RAM, and Windows 10. The results are measured with the highest precision in order to accurately validate the proposed model with minimum error as required for modern communication devices. The simulation was done for each task allocated to vehicles.

During the simulation, service execution in vehicular cloud networks was considered invariable. That means the relationship between each of the tasks was confirmed prior to the simulation commencement. Thus, each of the task degrees that are utilized to specify no. of children's tasks was persistent for the simulation procedure. Figure 10 shows the avg. service execution time of proposed algorithm and existing HEFT algorithm. As seen from this figure, execution time of proposed algorithm is reduced in comparison with the values of the existing H EFT algorithm that is set as limit for the services achievement by utilizing the proposed

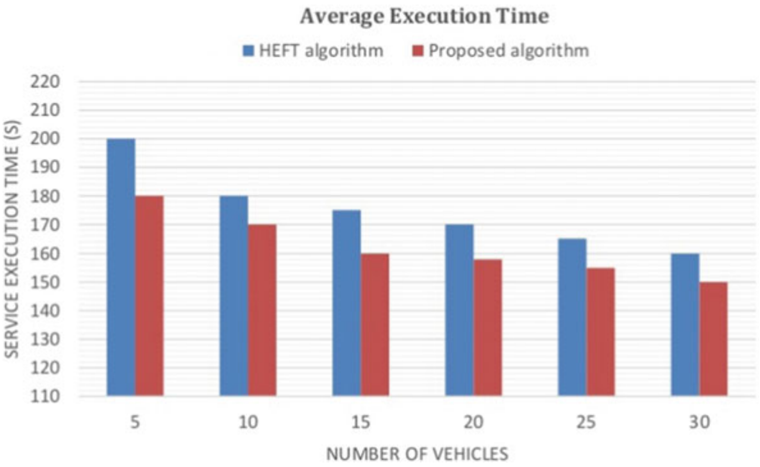


**Fig. 10** The average service execution time of proposed algorithm and existing HEFT algorithm

algorithm. Figure 11 depicts avg. service fails likelihood on the no. of the vehicles. It is seen that service fail likelihood of proposed algorithm is much lesser in comparison with the existing HEFT algorithm. The likelihood of existing HEFT algorithm is nearly 4.98% wherein the likelihood of proposed algorithm is nearly 1.9%. Figure 12 shows avg. service fail likelihood depending on the velocity of the vehicle in vehicle-to-infrastructure communication. Figure 13 illustrates avg. service fail likelihood depending on the velocity of the vehicle in vehicle-to-vehicle communication. Figures 12 and 13 clearly show that average service fail likelihood depending on the vehicle speed in the vehicle-to-infrastructure and vehicle-to-vehicle communication, respectively. The measured results demonstrate that proposed algorithm minimizes avg. time of execution for overall service in a pragmatic manner in comparison with the conventional Heterogeneous Earliest Finish Time (HEFT) algorithm.

## 4 Conclusion and Implications

The vehicles connectivity fault problem would lead to miscarriage of the services of the vehicular cloud networks (VCNs) in intelligent transportation systems (ITS) that is making the overall transportation very unsafe nowadays as the no. of vehicles are increasing rapidly across the globe. In order to resolve this huge problem, a novel model is presented in this paper for fault tolerance in VCN. The proposed algorithm for the validation of the suggested model is pragmatic in comparison with the traditional Heterogeneous Earliest Finish Time (HEFT) algorithm that is. safer as required for modern systems. The service fail likelihood of proposed algorithm is much lesser in comparison with the existing HEFT algorithm.

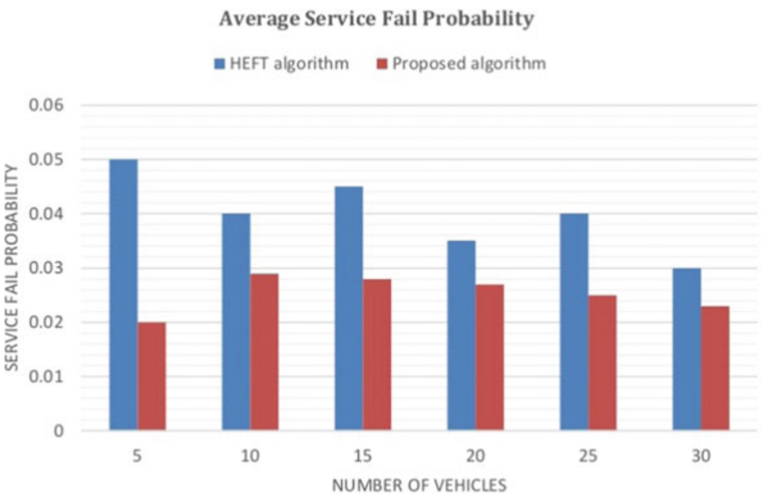


**Fig. 11** The average service fails likelihood on the no. of the vehicles

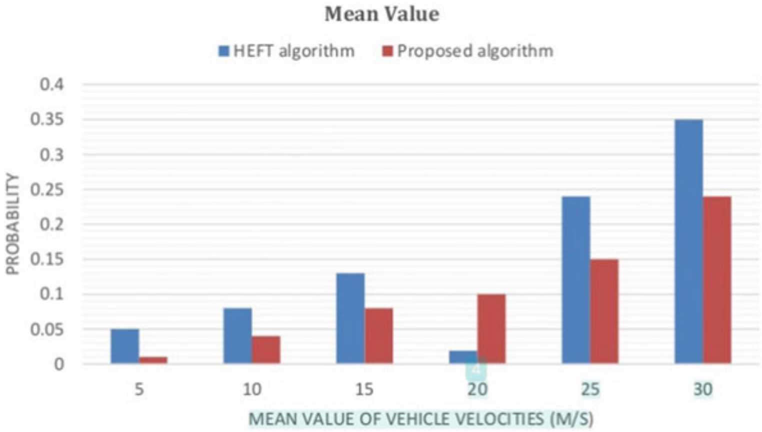


**Fig. 12** The average service fail likelihood depending on the velocity of the vehicle in vehicle-to-infrastructure communication

The likelihood of existing HEFT algorithm is nearly 4.98% wherein the likelihood of proposed algorithm is nearly 1.9%. The avg. service fail likelihood depending on the velocity of the vehicle in vehicle-to-infrastructure communication as well as for vehicle-to-vehicle communication is optimal and pragmatic. Moreover, several random parameters were taken for the simulation procedure for the optimal validation such as velocities, overall distance for the traveling, communication capability, workload, and many more. The measured results show that the proposed algorithm is pragmatic for the detection and reconfiguration of the node and link failure. Although,

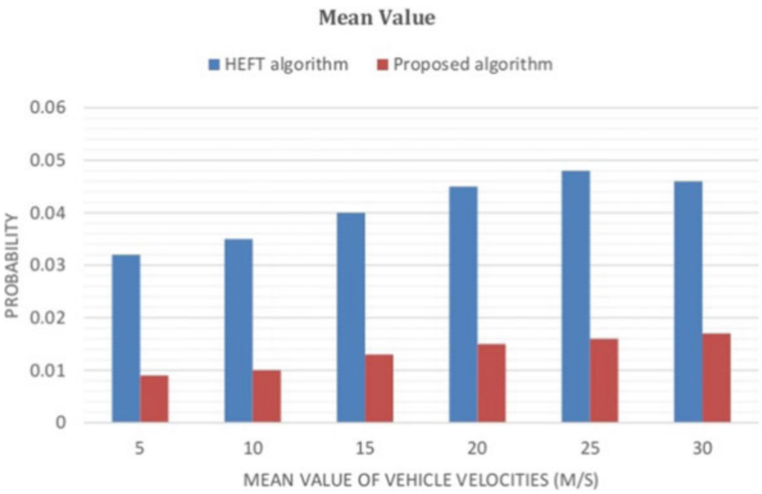


**Fig. 13** The average service fail likelihood depending on the velocity of the vehicle in vehicle-to-vehicle communication

extensive research has been carried out in this field during the last years, but there is enormous scope for more research in order to explore the full potential of this sector to find pragmatic solutions.

## References

1. Kakarla J, Siva Sathya S (2012)·A survey and qualitative analysis of multi-channel Mac protocols for VANET. Int J Comput Appl. http://doi.org/10.5120/4616-6844
2. Rammi B, Khaioun R, Fayad A, Khoukhi L (2017) Internet of things (IoT) technologies for smart cities. IET Netw
3. Park E, del Pobil AP, Kwon SJ (2018) The role of Internet of Things (IoT) in smart cities: technology roadmaporiented approaches. Sustainability.http://doi.org/10.3390/su10051388
4. Shretha R, Bajracharya R, Nam SY (2018) Challenge of future VANET and cloud-based approaches. Wirel Commun Mob Comput. http://doi.org/10.1155/201815603518
5. Deeksha, Kumar A, Bansal M (2017) A review on VANET security attack and their counter measure. http://doi.org/10.1109/ISPCC.2017.8269745
6. Hasrouny H, Samhat AE, Bassil C, Laouiti A (2017) VANet security challenges and solutions: a survey. Veh Commun. http://doi.org/10.1016/j.vehcom.2017.01.002
7. AJ-Heety OS, Zakaria Z, Ismail M, Shakir MM, Alani S, Alsariera H (2020) A comprehensive survey: benefits, services, recent works, challenges, security, and use cases for SDN-VANET. IEEE Access. http://doi.org/10.1109/ACCESS.2020.2992580
8. Dhankhar S, Agrawal S (2014) VANETs: a survey on routing protocols and issues. Int J Innov Res Sci Eng Technol 3(6):13427–13435
9. Newbegin M, Rengarajan A (2018) Improving the vehicular network (Vanet) QoS through various routing algorithm—a survey. Int J Eng Technol. http://doi.org/10.14419/ijet.v7i3.34.18961
10. Katiyar A, Singh D, Yadav RS (2020) State-of-the-an approach to clustering protocols in VAN ET: a survey. Wirel Netw. http://doi.org/10.1007/sl1276-020-02392-2

11. Li X, Li H (2014) A survey on data dissemination in VANET. Chin Sci Bull. http://doi.org/10.1007/s11434-014-0415-2
12. Alaya B, Khan R, Moulahi T, El Khediri S (2021) Study on QoS management for video streaming in vehicular ad hoc network (VANET). Wirel Pers Commun. http://doi.org/10.1007/s11277-021-08118-7
13. Hajlaoui R, Guyennet H, Moulahi T (2016) A survey on heuristic-based routing methods in vehicular ad-hoc network; technical challenges and future trends. IEEE Sens J. http://doi.org/10.1109/JSEN.2016.2583382
14. Huang J et al (2017) Synthesizing existing CSMA and TDMA based MAC protocols for VANETs. Sensors (Switzerland). http://doi.org/10.3390/s17020338
15. Khan UA, Lee SS (2019) Multi-layer problems and solutions in VANETs: a review. Electronics (Switzerland). http://doi.org/10.3390/electronics8020204
16. Dinesh MD (2014) Challenges in vehicle ad hoc network (VANET). Int J Eng Technol
17. Ur Rehman S, Khan MA, Zia TA, Zheng L (2013) Vehicular ad-hoc networks (VANETs)—an overview and challenges. J Wirel Netw Commun. http://doi.org/10.5923/jjwnc.20130303.02
18. Zeadally S, Hunt R, Chen YS, Irwin A, Hassan A (2012) Vehicular ad hoc networks (YANTIS); status, results, and challenges. Telecommun Syst. http://doi.org/10.1007/s11235-010-9400-5
19. Yang J, Fei Z (2013) Broadcasting with prediction and selective forwarding in vehicular networks. Int J Distrib Sens Netw. http://doi.org/10.1155/2013/309041
20. Chen W, Guha RK, Kwon TJ, Lee J, Hsu YY (2011) A survey and challenge in routing and data dissemination in vehicular ad hoc networks. Wirel Commun Mob Comput. http://doi.org/10.1002/wcm.862

# Multi-filter Enhanced Doctor and Patient Optimization Algorithm for Cancer Diagnosis

**Bibhuprasad Sahu, Saroja Kumar Rout, J. V. R. Ravindra, and Sachi Nandan Mohanty**

**Abstract** Medical data impacts a vital role in the diagnosis of cancer disease. Redundant and irrelevant featured genes are two different parameters that minimize the machine learning model's performance. Feature selection (FS) plays a crucial role in enhancing model performance. FS is used to attain good classification accuracy by reducing the overall computational cost of the machine learning model. In this paper, a three-stage filter wrapper model is proposed. The first phase consists of 4 different filters. An ensemble feature subset is generated using the output of the individual filter. The optimal feature subsets are input to the DOP algorithm. Experiments are carried out with six different benchmark datasets with various dimensions. The performance of the proposed model was evaluated with one of the well-known classifiers called the support vector machine. The empirical results indicate that the proposed model performed better than the accuracy of the individual filter with metaheuristic DOP-SVM.

**Keywords** Multi-filter · Microarray dataset · Cancer diagnosis · Doctor and patient optimization algorithm (DOP) · Feature selection

---

B. Sahu (✉)
Department of Artificial Intelligence and Data Science, Vardhaman College of Engineering (Autonomous), Hyderabad, India
e-mail: prasadnikhil176@gmail.com

S. K. Rout
Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, India

J. V. R. Ravindra
Department of Electronics and Communication Engineering, Vardhaman College of Engineering (Autonomous), Hyderabad, India
e-mail: jayanthi@ieee.org

S. N. Mohanty
Department of Computer Science and Engineering, Vardhaman College of Engineering (Autonomous), Hyderabad, India

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_7

## 1 Introduction

Feature selection (FS) is considered as a crucial stage of data pre-processing in the domains of machine learning (ML) and statistical data analysis. It is used to discover the subset of relevant features or attributes that boosts the model's performance [1, 2]. So the FS helps researcher to focus on the selected feature subset instead of deal with whole high-dimensional dataset which are normally high dimensional in nature with out losing the salient features. The technology enhancement not only helps the medical practitioners to recognize the stage of the disease; in other way, it generates large amount of data without any regularities. So the data cleaning plays a major role to recognize the disease [3]. While dealing with high-dimensional datasets, it take huge time to access each subset features. Three basic approaches of feature selection such as wrapper, filter, and embedded are preferred by the researcher to identify the efficient feature from the high dimensional datasets. In this current scenario, hybrid is the most preferable approach by the researchers [4, 5].

In filter approach, the classification model is not taken into consideration, it only focus to recognize the most featured subsets by rejecting the least statistical significant attributes. But reverse mechanism is adopted in wrapper. Here, the featured subsets are chosen according to the machine learning model. Depending on the evaluation criterion, some researchers use search operations with the wrapper. Embedded approach is quite similar to the wrapper, but here, the feature selection is performed at the same time of model training [6]. As the classifier is dataset dependent, so in some case, it may not provide the expected result. Hybrid approaches are the combo operation of both filter and wrapper. Hybrid approaches provide most accurate and less execution time as compared to wrapper, as each and every machine learning algorithms have own advantages and disadvantages, it is the important task to chose the suitable one keeping eye on the weakness also.

After reviewing the characteristics of different feature selection approaches, we have proposed a 3-stage FS model for our work. We have considered 4 filter approaches in the first phase, correlation filter is used in the second phase, and DOP algorithm is used as wrapper to detect the featured subset genes. Individual filter is applied to each dataset, and the union of all filter is taken into consideration to prepare the common one for evaluation with classifier like SVM. The main motto behind the multi-filter FS is, as filter having own evaluation criteria, if we consider one filter, then there may be a chance that important feature may left. So in this study, we focus multiple filters to provide a opportunity that none of an important feature should miss, which may degrade the accuracy proposed model. And using correlation filter is used to remove low correlated features, and finally, it may considered as input for the doctor patient optimization algorithm (DOP) to achieve high accuracy.

For our experimental study, we have taken high-dimensional microarray datasets such as leukemia [7], SRBCT, Tumors_9 [8], breast cancer [9], colon cancer, and lung cancer [10]. Datasets for this study are downloaded from www.gems-system.org and UCI repository https://archive.ics.uci.edu. While dealing with medical datasets, it is important to select the optimal feature subset out of high dimension which provide the better opportunity to accurate and easy diagnosis of the disease and less no's of medical tests may require.

**Original contribution of the paper**

1. Three stage wrapper filter wrapper model is designed. Where the 1st phase consists of 4 different filters. Ensemble feature subset is generated taking the output of the individual filter into the consideration.
2. In the 2nd phase, the optimal feature subsets are inputted metaheuristic DOP algorithm.
3. In 3rd phase, classifier like support vector machine (SVM) is used to study the non-bias performance of the proposed model.
4. After rigorous literature survey, we found that various researchers used multi-filter (MF) feature selection approach for different wrapper algorithms, but this will be the first ever research done with DOP algorithm with cancer datasets.

The flow of work progress is follows: The motivation and existing related work are presented in Sect. 2, and proposed model description is briefed in Sects. 3 and 4 contain the details of the datasets used in this experimental purpose along with result analysis. Section 5 concludes the study with future scope.

## 2 Background Study

FS plays a important role to enhance the performance of the model by deploying various metaheuristic approaches. In this section, it presents the different novel approaches used by various researchers according to the nature of the dataset.

A 2-tier model is proposed by Deng et al. [11], in which XGBoost algorithm is used to rank the feature genes from the dataset during the first phase. Then, the author deployed a multi-objective GA (MOGA) which is used to identify the optimal features. The experimental results prove that the proposed model able to detect 54 no's of optimal genes with 98% accuracy using SVM classifier. The author in [12] proposed a hybrid approach with a combination of PCA-SVM. This model achieved an accuracy of 94.17%. A multi-tier filter with wrapper model [13] is developed to detect the feature subsets from the high-dimensional datasets. Out of 4 filters and classifiers, SNR-WOA combination achieves the accuracy of 99.48% with ANN classifier. The author [14] proposed a hybrid PCA-ANN-GA based model, in which PCA is used for dimension reduction by identifying best 10 highest eigenvalues. Then, ANN-GA is used to select the optimal genes. Salem et al. [15] presented a hybrid IG-GA-GP model for classification of microarray datasets. In [16], MS-FS is used as filter and high ranked features taken into consideration for simple voting scheme for classification.

We reviewed similar existing hybrid models of different researchers in the literature survey for detection optimal feature subset from the microarray datasets. We have motivated to propose a novel frame work based on multi-tier architecture along with multiple filters and wrapper method to recognize the most significant genes.

### 2.1 Filter-Wrapper-Based FS

The main concept behind the filter methods is divided into 2 parts such as ranking and evaluation. Ranking process can be performed using different heuristic such as joint mutual information (JMI), maximum relevance minimum redundancy (MRMR), ReilefF, and information gain (IG). The subset features are selected based on ranking concept. Still, there is not consistent acceptance that the selected feature subset will provide best solution. So the performance of the can be tested with a wrapper method.

### 2.2 Doctor and Patient Optimization Algorithm

Doctor and patient optimization algorithm was proposed by Dehghani, Mohammad, et al. on the year of 2020 based on the treatment procedure of a doctor toward the patient [17]. The overall procedure patient is divided into three basic phases such as (1) vaccination, (2) drugs assignment, and (3) surgery. In this optimization technique, the phase 1 vaccination is consider as population which is used for preventing the disease. Then, phase 2 and 3 deal with the drug assignment to the patient as per the desired disease and if need surgery may be taken into consideration, respectively. The population for patient ($P$) consists of $i$th no of patients, where $p_i^d$ presents $d$th feature of the $i$th patient. Suppose there are $N$ no of patients in the population. In this algorithm, population is evaluated in three phases such as mentioned above. For processing for all three phases, below equations are taken into consideration.

$$\text{dose}_i = 2 - \frac{\text{fit}_i^n}{\text{fit}_{\text{best}}^n} \tag{1}$$

$$\text{fit}_i^n = \frac{\text{fit}_i - \text{fit}_{\text{worst}}}{\sum_{j=1}^{N} (\text{fit}_j - f_{\text{worst}})} \tag{2}$$

$$f_{\text{worst}} = \max(\text{fit}) \text{ and } \quad P_{\text{worst}} = P(\text{location}(f_{\text{worst}})) \tag{3}$$

$$f_{\text{best}} = \min(\text{fit}) \text{ and } \quad P_{\text{best}} = P(\text{location}(f_{\text{best}})) \tag{4}$$

Equation (1) presents the dose provided to the $i$th patient. Normalized fitness value of the $i$th patient is presented in Eq. (2). Equations (3) and (4) present the fitness function of worst ($f_{\text{worst}}$) and best patient ($f_{\text{best}}$) and position of worst ($P_{\text{worst}}$) and best ($P_{\text{best}}$) patient, respectively.

In the vaccination phase, the patients in the community are vaccinated to avoid the disease in near future. The simulation of the this phase presented in Eqs. (5) and (6).

$$\text{vac}_i^d = \text{rand} \times (\text{dose}_i \times p_i^d - p_{\text{worst}}^d) \tag{5}$$

$$\mathrm{vac}_i^d = \mathrm{rand} \times (\mathrm{dose}_i \times p_i^d - p_{\mathrm{best}}^d) \tag{6}$$

Here, the *d*th dimension of the *i*th patient for the vaccination process and $p_{\mathrm{worst}}^d$ and $p_{\mathrm{best}}^d$ presents the *d*th dimension of worst and patient respectively. And rand is a random value which lies between $0 - 1$.

In the drug assignment phase, the doctor prescribes the drug to the patient as per requirement of the disease. This is done as per the previous condition of the patient. The overall simulation is presented in the below Eqs. (7) and (8).

$$d_i^d = \mathrm{rand} \times (p_{\mathrm{best}}^d - \mathrm{dose}_i \times p_i^d) \tag{7}$$

$$P_i = \begin{cases} P_i + d_i, & \text{if } \mathrm{fit}(P_i + d_i) \le \mathrm{fit}_i. \\ P_i, & \text{otherwise.} \end{cases} \tag{8}$$

Here, $d_i^d$ is the *d*th dimension of *i*th patient, and $p_{\mathrm{best}}^d$ is the *d*th dimension of best patients.

Surgery is the last option preferred by the doctor when first two phases are not sufficient or in case of any serious conditions. With surgery, the condition of the patient grows with surgery. This simulation of the phase 3 is mentioned in Eq. (9).

$$P_i = \begin{cases} 0.6 \times P_i + 0.4 \times P_{\mathrm{best}}, & \text{if } F_{\mathrm{best}}^n \ge 0.9 F_{\mathrm{best}}^n \\ P_i, & \text{otherwise.} \end{cases} \tag{9}$$

The pseudo code of MFS-DOP is stated as below.

**Algorithm 1:** Doctor and Patient Optimization Algorithm.

```
Input the filtered feature subset out of 4 filters using rank aggregation method.
Begin DOP algorithm.
Initialization of patient population (P).
for Every time step 1: max iteration do
    Evaluation of fitness function is done.
    Using equation 3 and 4 update $f_{worst}$,$f_{best}$,$P_{best}$ and $P_{worst}$.
    Using equation 2 update the fitness value ($fit_i^n$).
    for for all i=1:N do
        Using equation 1 update $dose_i$.
        Update the value of $P_i$ in all three phases using equation 7,8,9.
loop end
    Save the value of $f_{best}$ and $P_{best}$.
Iteration loop end
Save the best solution
End of the algorithm
```

### 2.3 Ensemble Techniques for Individual Filter Output

In this study, four different filters are used to identify the feature subset described in Sect. 3. As different filters use their own statistical method to detect the feature genes, and we have to prepare an optimal feature subset with top *N* features which are significant in nature.

Let us consider the nos of datasets was *D*, nos of filters used in this study was *F*. The gene set *D*, which contained the top *N* genes in *L* datasets, was obtained as follows (where *S* was assumed to be more than *N*):

1. Suppose $G = g_1, g_2, \ldots, g_s$ and $D' = \varphi$.
2. Use the filter $F_i$ to calculate the statistical value and rank them, where $i \in i = 1, 2, \ldots, F$.
3. Select the *N* top ranked features of individual filter and add it to *D* and delete the *N* genes from *G*.
4. Prepare the union of the *F* filters genes by removing the duplicate and prepare the final dataset $D'$.

## 3 Proposed Model

Figure 1 presents the workflow of the proposed model. The proposed model consists of 3 phases/tiers. In the first tier, it presents the multi-filter ensemble feature selection model which consists of 4 filter approaches using rank aggregation to evaluate the high ranked feature subsets. In tier 2, the evaluated high ranked feature subset is feed as input to the DOP algorithm to identify the optimal feature subset. Later in tier 3, the performance of the MFS-DOP is evaluated using classifier such as SVM (Fig. 2 and Tables 1, 2).

## 4 Experimental Study

The proposed MFS-DOP algorithm was implemented on the bench marked microarray cancer datasets. The overall experimental analysis is done through MATLAB 2013(b), 8 GB RAM with Windows 10 operating system. As per the Sect. 2.3, the MF-FS is done using Java. Then, a top ranked dataset is prepared as input for the new metaheuristic algorithm called DOP. In this study, we have compared the performance of the proposed model which is compared with individual filter. The accuracy of the different datasets is performed with different range of gene subset such as 10, 20, 30, 40, and 50. The range of the accuracy achieved while tested with the individual filter is 71–99%. For SRBCT dataset, MFS-DOP-SVM achieved the better accuracy of 98.36%. In case of Tumors9 and Leukaemia3, the accuracy achieved by the proposed methodology is 97.98% and 92.63%, whereas 99.44%, 97.81%, and

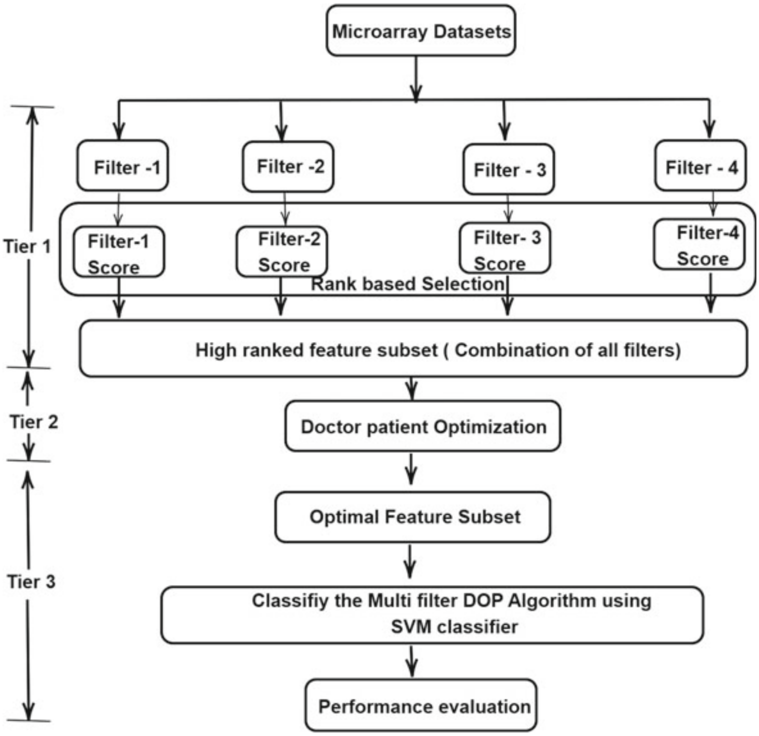


**Fig. 1** Block diagram of the proposed model MFS-DOP-SVM

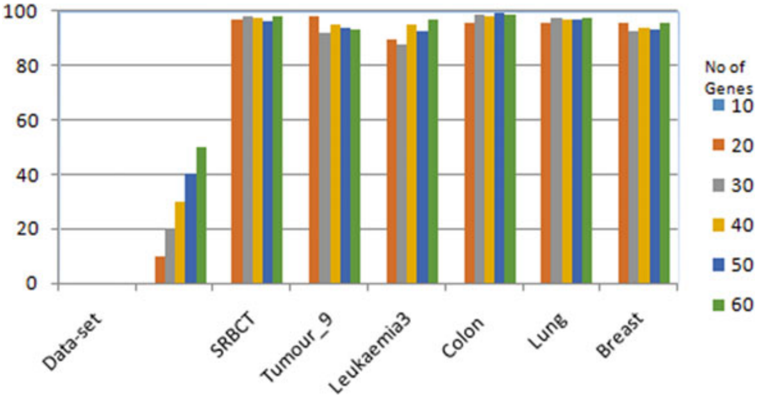


**Fig. 2** Performance of proposed model with SVM classifier

**Table 1** Accuracy evaluation with SVM classifier

| Dataset | Method | Number of gene subsets | | | | |
|---|---|---|---|---|---|---|
| | | 10 | 20 | 30 | 40 | 50 |
| SRBCT | Relief-F-DOP-SVM | 92.52 | 96.44 | 96.56 | 96.19 | 96.52 |
| | IG-DOP-SVM | 88.84 | 79.66 | 89.36 | 83.08 | 88.07 |
| | mRMR-DOP-SVM | 87.94 | 88.01 | 84.97 | 86.34 | 92.36 |
| | JMI-DOP-SVM | 92.19 | 91.81 | 90.91 | 96.39 | 95.37 |
| | MFS-DOP-SVM | 97.06 | 98.36 | 97.69 | 96.34 | 97.99 |
| Tumors_9 | Relief-F-DOP-SVM | 88.94 | 89.67 | 90.68 | 94.89 | 93.78 |
| | IG-DOP-SVM | 91.69 | 90.89 | 88.67 | 87.09 | 84.97 |
| | mRMR-DOP-SVM | 94.52 | 93.12 | 95.05 | 97.81 | 96.17 |
| | JMI-DOP-SVM | 90.68 | 94.89 | 93.78 | 89.67 | 88.94 |
| | MFS-DOP-SVM | 97.98 | 96.89 | 95.37 | 94.01 | 93.03 |
| Leukaemia3 | Relief-F-DOP-SVM | 86.83 | 84.34 | 85.65 | 81.39 | 82.96 |
| | IG-DOP-SVM | 74.07 | 75.06 | 72.98 | 79.13 | 82.36 |
| | mRMR-DOP-SVM | 80.36 | 82.93 | 84.59 | 86.16 | 87.96 |
| | JMI-DOP-SVM | 83.96 | 91.39 | 88.18 | 83.66 | 89.03 |
| | MFS-DOP-SVM | 89.34 | 87.61 | 94.85 | 92.63 | 96.94 |
| Colon | Relief-F-DOP-SVM | 97.44 | 98.89 | 96.84 | 97.46 | 97.91 |
| | IG-DOP-SVM | 91.31 | 90.49 | 90.39 | 89.89 | 94.17 |
| | mRMR-DOP-SVM | 82.66 | 86.33 | 88.46 | 89.06 | 88.39 |
| | JMI-DOP-SVM | 81.75 | 79.09 | 74.54 | 78.26 | 77.08 |
| | MFS-DOP-SVM | 95.83 | 98.66 | 97.94 | 99.44 | 98.66 |
| Lung | Relief-F-DOP-SVM | 87.03 | 96.39 | 94.59 | 98.65 | 88.03 |
| | IG-DOP-SVM | 97.73 | 98.61 | 97.72 | 99.81 | 89.46 |
| | mRMR-DOP-SVM | 97.84 | 98.93 | 96.88 | 99.18 | 98.21 |
| | JMI-DOP-SVM | 96.39 | 91.09 | 96.17 | 93.39 | 90.87 |
| | MFS-DOP-SVM | 95.63 | 97.81 | 97.04 | 96.97 | 97.81 |
| Breast | Relief-F-DOP-SVM | 93.23 | 80.06 | 93.39 | 90.81 | 88.69 |
| | IG-DOP-SVM | 99.81 | 89.46 | 94.17 | 93.64 | 97.81 |
| | mRMR-DOP-SVM | 81.06 | 89.39 | 93.39 | 87.96 | 94.39 |
| | JMI-DOP-SVM | 99.18 | 98.21 | 87.98 | 97.02 | 93.98 |
| | MFS-DOP-SVM | 95.89 | 92.37 | 94.04 | 93.19 | 95.81 |

95.89% in case of colon, lung, and breast cancer, respectively. Even if the results are not so good, but it is quite impressive in comparison with individual filter combine with DOP-SVM.

As per the literature survey,in comparing with existing approaches, the result of the proposed model is not good. But it is quite close to expected one. Due to page limit constraint, we unable to share the comparison table. This is the first ever study of the DOP with cancer dataset. Addition of different approaches or tuning the different parameters of the SVM may enhance the performance of the model [18–20].

**Table 2** Accuracy evaluation with SVM classifier

| Dataset | TLBO-SA | IG-MBKH | R-ACO | BSFLA-PSO | ALO | R-SMO-SA-SVM |
|---|---|---|---|---|---|---|
| $DS_1$ | 95.31 | – | 95.06 | – | 94.79 | **96.79** |
| $DS_2$ | 96.98 | – | – | – | 87.41 | **98.11** |
| $DS_3$ | – | – | – | – | 88.17 | **95.91** |
| $DS_4$ | **99.87** | – | 99.50 | 94.91 | 89.57 | 99.21 |
| $DS_5$ | **99.01** | 96.47 | 94.00 | – | – | 96.98 |
| $DS_6$ | 95.31 | **100** | 95.80 | 95.78 | 87.39 | 98.45 |
| $DS_7$ | 99.54 | **100** | – | – | 91.44 | 96.78 |
| $DS_8$ | 99.91 | **100** | – | – | – | 94.31 |
| $DS_9$ | **99.52** | – | – | – | – | 99.45 |
| $DS_{10}$ | **99.13** | – | 89.20 | 96.76 | 89.44 | 99.11 |

## 5 Conclusion

A first ever MFS doctor patient optimization algorithm is used to identify the informative genes from the high-dimension microarray benchmark datasets. The overall experiments are carried out with 6 different datasets. To obtain high quality result, DOP algorithm is used based on the concept of patient treatment assignment concept. From the experimental study, we found that implementation of ensemble filter algorithm enhances the feature selection process in comparison with single filter. The performance of this model is not better than the existing models but quite efficient and accurate. The improvement can be achieved with tuning various parameters of SVM and DOP which may be taken into consideration.

## References

1. Ghosh M, Guha R, Singh PK, Bhateja V, Sarkar R (2019) A histogram based fuzzy ensemble technique for feature selection. Evol Intell 12(4):713–724
2. Ghosh KK, Ahmed S, Singh PK, Geem ZW, Sarkar R (2020) Improved binary sailfish optimizer based on adaptive $\beta$-hill climbing for feature selection. IEEE Access 8:83548–83560
3. Duval B, Hao JK, Hernandez JC (2009) A memetic algorithm for gene selection and molecular classification of cancer. In: Proceedings of the 11th annual conference on genetic and evolutionary computation, pp 201–208
4. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Comput Electr Eng 40(1):16–28
5. Ye H, Gao J, Ding Y (2007) A generalized Gronwall inequality and its application to a fractional differential equation. J Math Anal Appl 328(2):1075–1081
6. Lu H, Chen J, Yan K, Jin Q, Xue Y, Gao Z (2017) A hybrid feature selection algorithm for gene expression data classification. Neurocomputing 256:56–62
7. Mandal M, Singh PK, Ijaz MF, Shafi J, Sarkar R (2021) A tri-stage wrapper-filter feature selection framework for disease classification. Sensors 21(16):5571

8. Filippini T, Heck JE, Malagoli C, Giovane CD, Vinceti M (2015) A review and meta-analysis of outdoor air pollution and risk of childhood leukemia. J Environ Sci Health Part C 33(1):36–66
9. Sahu B, Panigrahi A, Sukla S, Biswal BB (2020) MRMR-BAT-HS: a clinical decision support system for cancer diagnosis. Leukemia 7129(73):48
10. Sahu B (2018) A combo feature selection method (filter + wrapper) for microarray gene classification. Int J Pure Appl Math 118(16):389–401
11. Deng X, Li M, Deng S, Wang L (2021) Hybrid gene selection approach using XGBoost and multi-objective genetic algorithm for cancer classification. arXiv preprint arXiv:2106.05841
12. Sahu B, Mohanty S, Rout S (2019) A hybrid approach for breast cancer classification and diagnosis. EAI Endorsed Trans Scalable Inf Syst 6(20)
13. Prabhakar SK, Lee SW (2020) Transformation based tri-level feature selection approach using wavelets and swarm computing for prostate cancer classification. IEEE Access 8:127462–127476
14. Cahyaningrum K, Astuti W (2020) Microarray gene expression classification for cancer detection using artificial neural networks and genetic algorithm hybrid intelligence. In: 2020 international conference on data science and its applications (ICoDSA). IEEE, pp 1–7
15. Salem H, Attiya G, El-Fishawy N (2017) Classification of human cancer diseases by gene expression profiles. Appl Soft Comput 50:124–134
16. Bolón-Canedo V, Sánchez-Maroño N, Alonso-Betanzos A (2012) An ensemble of filters and classifiers for microarray data classification. Pattern Recogn 45(1):531–539
17. Dehghani M, Mardaneh M, Guerrero JM, Malik OP, Ramirez-Mendoza RA, Matas J et al. (2020) A new “Doctor and patient” optimization algorithm: an application to energy commitment problem. Appl Sci 10(17):5791
18. Sahu B, Mohanty SN (2021) CMBA-SVM: a clinical approach for Parkinson disease diagnosis. Int J Inf Technol 13(2):647–655
19. Sahu B, Badajena JC, Panigrahi A, Rout C, Sethi S (2020) 7 an intelligence-based health biomarker identification system using microarray analysis. Appl Intell Decis Making Mach Learn 137–161
20. Sahu B (2019) Multi-tier hybrid feature selection by combining filter and wrapper for subset feature selection in cancer classification. Indian J Sci Technol 12(3):1–11

# Disease Diagnostic Model Using Fine-Tuned Intensive Learning for Medical Image

**Prasanalakshmi Balaji, Prabha Selvaraj, Vijay Kumar Burugari, and Kanmani Palaniappan**

**Abstract** Computer-Aided Diagnostic (CAD) techniques use image-based categorization as an effective method. Existing approaches rely primarily on structure, coloring, texture-based, and mixtures. The majority of these is issue-specific and demonstrate complementary in medical images, resulting in a scheme that cannot represent high-level problem functional requirements and has impoverished prototype generalization capability. Modern deep neural networks have made it possible to build an edge model which can generate the best classification categories from primary healthcare image data. The Disease Diagnostic Model utilizing Fine-Tuned Intensive Learning (DDM-FTIL) classification offers superior characteristics generated from a convolution neural network (CNN) with some classical characteristics. Adaptive mean filtering and adaptive histogram are used to strengthen and optimize the pixel density. We use the Gabor filter to convert the improved images. The characteristics are retrieved by the Gray-Level Co-occurrence Matrix (GLCM) as well as Local Binary Pattern (LBP) descriptors before classing the DDM-FTIL classifier. Oral cancer areas were segmented using morphometric approaches.

**Keywords** Convolutional neural network · Deep learning · Image processing · Feature extraction

---

P. Balaji (✉)
Department of Computer Science, King Khalid University, Abha, Saudi Arabia
e-mail: prengaraj@kku.edu.sa; drsanaksa@gmail.com

P. Selvaraj
School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
e-mail: prabha.s@vitap.ac.in

V. K. Burugari
Department of CSE, Koneru Lakshmaiah Education Foundation, K L University, Vaddeswaram, Andhra Pradesh, India

K. Palaniappan
Department of DSBS, SRM Institute of Science and Technology, Chennai, Tamilnadu, India

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_8

## 1 Introduction

Mouth cancer is the most common are among the most life-threatening diseases. Cancer disease kills one person out of every six. According to the WHO, cancer is the second largest cause of illness and death worldwide. According to the research, mouth cancer patients die from oropharyngeal cancer. Oral cancer, often known as mouth cancer, affects the lips, jaw, tongues, and neck. Mouth cancers come from cigarettes, drug use, massive alcohol drinking, and a hereditary type of disease. Improved therapies like radiotherapy and chemotherapeutic have now become accessible. However, the mouth cavity must pick early to preserve a person's life. For soft tissue prognostic, the MRI-scan method appears to be ideal. Use many classification algorithms to separate the areas of aberrant images.

Feature extraction, binarization, sector growth, and morphology analysis are the main approaches to image separation. Clusters, model evaluation, and probabilistic approaches are sophisticated segmentation techniques. The literature includes seven different techniques for the method of image separation including a segmentation method that depends on the edges, in which the shift in pixel intensity of the image plays an important role. It finds the image's edges and connects them to the bounds needed to split the relevant areas. It performs effectively with photos that have sufficient contrast. However, it is not suitable for a large number of edges. The thresholding technique divides the images based on the threshold value. A histogram is a visual depiction with different intensities. Therefore, it will not necessitate the use of any primary variables. The disadvantage of this method is that it is dependent on the maxima. A basic approach to clustering the images based on specific criteria of location is another method used for image separation. A zone is a collection of dots having comparable qualities in an image. It is even more noise resistant and performs much with comparable requirements, although time and space-consuming. Clustering of images can also be done based on some standard criteria. The information pieces are grouped depending on how comparable they are. It comes in handy in practical systems. For example, it is not easy to sort out the method to determine the degree of membership. Watershed algorithms employ geometric inferences, in which the intensity depicts the valley, and each zone has a water flow. As a result, the findings are much more consistent, and the borders discovered are unbroken. However, it necessitates complicated gradient computations. As a result, we use the DDM-FTIL classification throughout the work. Even techniques that rely on partial differential equations are the most efficient solution for interval-based tasks is to implement the dataset using differential calculus. The downside to this approach is that it is computationally complicated. Finally, the techniques based on neural networks are used. A neural network is a set of neural connections linked by their weight. For active learning, the artificial neural network replicates brain activity. It is perfect for healthcare photos.

The discussed methods of image separation also depend on the accuracy of the imaging technique. In the realm of healthcare, digital imagery is hugely significant. Modern medicine continuously evolves as new imaging modalities such as image pre-processing, augmentation, and evaluation become available. The ability to split images in identifying contaminated areas is an essential feature of diagnostic imaging. We utilize the imaging produced by an X-ray, C-scan, ultrasonic, MRI, and PET to diagnose and identify various illnesses and make high standard vision. X-rays are electromagnetic fields are that captivate the body's natural interior components. The radiation dose received is insignificant. This is because the frequency of soft X-rays is only 10 nm in size. The computerized tomography incorporates information from many X-rays to provide a complete version of the body's architecture. It delivers a three-dimensional cross-sectional image of the organism. Magnetic Resonance Imaging is a powerful magnetic field that captures images of interior parts with radioactivity radiation. Unlike CT-scan and PET, it employs radiation exposure. However, it does have a great deal of spatial and temporal resolution. As a result, they seem to be very effective at spotting soft tissue cancer cells, tendons, and bone fragments. Positron Emission Tomography is a nuclear reaction scanning that employs specific dyes that incorporate radioisotopes. Such detectors aid in the detection of illness in the chest and abdomen.

All images were downloaded in JPG format from the Kaggle database during the first stage [1]. After that, we did color normalization before feeding the dataset into the deep neural network. We used an available online collection of images of the mouth and tongue for the disease identification system. These classifications were discovered in the cancer records. We only created two sections of a vital organ, the lips, and tongue, enabling cancer detection considering the time restrictions.

## 2 Related Works

Through an assessment of long-term results, Youssef et al. [2] employed the Face and Mouth Tumor Track to diagnose, cure, maintain, and detect mouth cancer. It employs a revolutionary feature selection algorithm known as SVM, based on a hierarchical model, to identify the best characteristics in a group of photos. Chan et al. [3] integrated deep neural networks and textural maps. In a single model, the ROI was used to designate the malignant spots found dynamically. The standard error values of the textured image are computed using a sliding window. These values are employed to develop a textured map divided into several patches and supplied into the deep convolutional model as data input. The morphology watersheds approach was developed by Anuradha et al. [4]. To categorize aberrant oral MRI images and locate and segment cancerous areas, watersheds, and markers using watershed segments techniques. The authors used a watershed segmentation method to solve 85.2% segmentation performance and a marker-regulated watershed segmentation method to achieve 90.25% segmentation results. Ghosh [5] suggested a unique Neuro-fuzzy categorization approach for data analysis. A generic bell-shaped association function was used to fuzzify the inputs. It made use of a fuzzification matrix, which assigned a degree of association to each group or class. Jung et al. [6] employed an optical coherence tomography approach for detecting and characterizing cancer areas in mouth MRI images. The researchers scattered the original mouth MRI image to improve the classification performance and segmentation accuracy.

In addition, the researchers assessed the effects of linear and semi-categorization techniques. Ram et al. [7] suggested a computer-assisted method for detecting and classifying malignant areas in MRI images of the mouth. The scientists categorized the original mouth MRI-scan either as healthy or malignancy epithelium diseases based on the percentage of cancer-infected areas in mouth MRI images. Zare et al. [8] created an ongoing classification scheme that generates four classifier models using distinct categories to improve accuracy—making Healthcare images at an increased pace due to fast improvements in healthcare technology. Physically watching these photos is too expensive, and it is prone to errors. Consequently, an accurate categorization employs information directly collected from visuals. Before merging the photos into a unified platform, one popular way is to break them into divisions and independently obtain visual information from each section of a location. Xu et al. [9] create a method that completes elevated jobs like detection and segmentation with little labeled training data and excellent features visualizations. This research investigates the effectiveness of finishing elevated discussions with minimum summarizing and outstanding features analyses for healthcare images. In healthcare image segmentation, visuals like organelles have primary diagnostic functions. SIFT and HARR, two previously discovered features, we cannot fully characterize these atoms. As a consequence, the importance of features extracted cannot overstates. Wan and Wang [10] suggest that effective feature descriptions and closeness measurements are crucial to the extract effectiveness of a content-based image retrieval (CBIR) technology. Despite centuries of study, it remains one of the most challenging open challenges that significantly inhibits the development of accurate CBIR systems. The most severe problem has been identified as a semantics gap among pixels from a low-resolution image by computers and formal semantics evaluated by people. By exploring deep networks to gain knowledge characteristics at various levels of complexity from necessary data, deep learning techniques facilitate a scheme. To realize complex jobs that substantively trace basic tactile model parameters of outcome without focusing on pertinent domain expertise. Krizhevsky [11] et al. Present object recognition algorithms use machine learning considerably. We can enhance their positions by gathering more samples, developing more effective methodologies, and utilizing the most potent anti-techniques. They recently obtained previous studies containing tagged photos numbered hundreds of thousands. Simple identification activities could be solved effectively with datasets of this length, especially by adding tag transformations. Li et al. [12] developed a tailored deep neural network to categorize lung image patches. This prototype used only one convolution layer to conquer the generalization error and retrieve in-depth features. It delivers the highest categorization efficiency concerning SIFT features, local binary characteristics, and unsubstantiated pattern recognition using a restricted Boltzmann machine. The researchers developed a basic deep neural network model called principle component assessment networking (PCANet) that has been employed by Zeng et al. [13] paired with the spatial transmission data of color images to attain province classification performance in multiple datasets [14]. The authors used a CNN learned by ImageNet to describe specific sorts of diseases in chest X-ray images [15]. By mixing features retrieved from the CNN with handmade characteristics, they could attain the most remarkable accuracy results. Shin et al. [16] talked about how transfer learning may help with medical imagery. They also demonstrated their findings in identifying thoracoabdominal lymph nodes (LNs) and categorizing lung infections. For lung cancer identification, Rakotomamonjy et al. [17] used the dispersion change, which was initially proposed by Bruna and Mallat [18], to mine features paired with LBP and LQP, which were demonstrated to be resilient to tiny displacements in the images. They also tested the results and efficacy on the 2D-Hela and Pap smear datasets. Cruzroa et al. [19] proposed a deep learning strategy for automatically detecting invasive ductal carcinoma (IDC) tissues areas in whole slide images (WSI) of cancer, which was validated using data of 162 diagnosed cases with IDC, with a balancing performance of 84.23.

## 3 Proposed System

The suggested technique identifies and splits the locations of mouth cancer depicted in Fig. 1. The stages in the procedure are as shown in Fig. 2: Pre-processing comes first, followed by image enhancement. Finally, categorize the image using Gabor Transforming and Feature Extraction.

Next, the distortion in the source photos is removed using an adaptive mean filter, the processed image is improved using an adaptable histogram-based method. Next, the image's outcome is changed using the Gabor Transformation, yielding a multi-resolution image. Finally, a DDM-FTIL classifier is used to extract features from the resulting image then categorize them.

**Image Preparation**. Numerous surrounding disturbances, movements seen between the sensor and the target, missing focus, and inappropriate shutters release affect photos throughout the procurement procedure. The disturbance determines the image performance in the image pixel. Therefore, the pixel value must be de-noised to increase the image's clarity as shown in Fig. 3. While de-noising the filters, the borders of the image's impact when using traditional filtering. An adaptive technique is employed to modify the values of the digital filter to reduce noise. The noise removal approach does not affect the image's boundaries when used on an image. As a result, the adaptive median filter overcomes the limitations of the traditional technique. The adaptive median filter is used for the image pixels to reduce noise and improve image resolution.

**Improvement of the image**. The photos captured with the camera systems are of limited resolution. The adaptable histogram-based approach boosts each pixel's grayscale luminosity. This procedure applies to the processed images' outcome. The stages used in the histogram equalization method are as follows:

**Fig. 1** Sample image data set

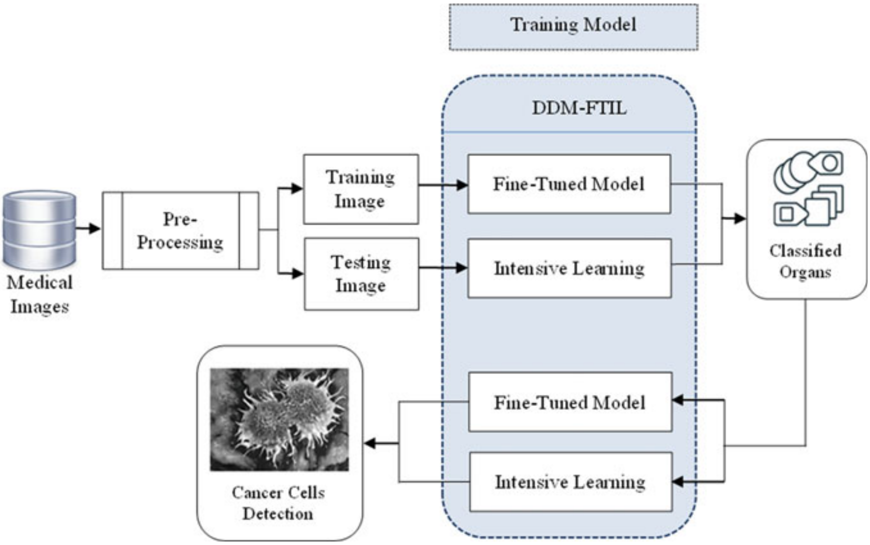


**Fig. 2** Proposed DDM-FTIL architecture

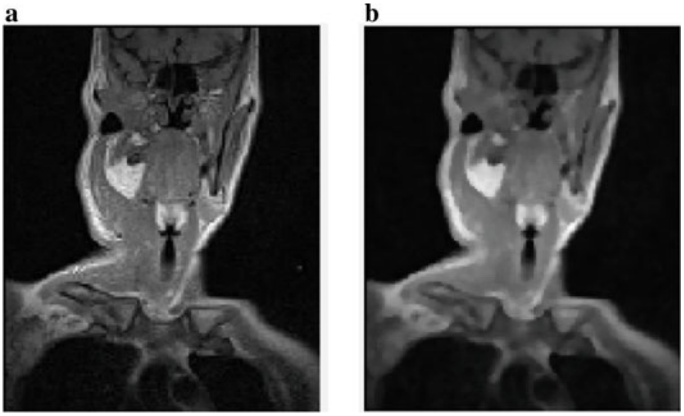


**Fig. 3** **a** Original image, **b** de-noised image

1. Separate the processed photos into 8 × 8 non-overlapping submodules.
2. The next aim is to define each subblock's histogram frequency.
3. On each histogram statistic, find the summary statistics.

In the mouth cancer images, swap the pixels with cumulative probability statistics. When the conclusion of the pixels in the mouth cancer image is achieved, cease as in Fig. 4.

**Gabor Image transformation (GIT)**. The GIT transforms the spatial dimension into a multi-pixel image that links the frequency response of each pixel. There are tangible and intangible components in the Gabor amplified image. The Gabor amplitude is calculated by choosing the highest image pixels at each place in the image as in Fig. 5.

**Extraction of Characteristics**. An image's characteristics use to differentiate one item from another. The different characteristics that represent particular attributes

**Fig. 4** Improvised image

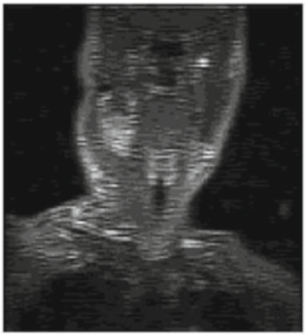


**Fig. 5** Gabor magnitude oral image

of an item are known as unusual characteristics. The distinctive characteristics employed in this study are the Local Binary Pattern (LBP) and Gray-Level Co-Occurrence Matrix (GLCM).

- Binary Patterns in a Local Region

It is a texture descriptor [20] used to categorize an image that's been Gabor altered. The LBP performs in the following manner:

Choose neighboring pixels for each pixel with a radius in an image. Calculate the brightness gap between the observed pixels and their surroundings [21]. Create a bit vector for increased proportion and zeroes for unfavorable disparities. Transform the bit vectors and intensity value to their decimals equivalents. The recovered LBP characteristic image and the Gabor magnitude mouth image seem to be the same size, and the recovered LBP characteristic image shows in Fig. 6.

- GLCM

It is a quantitative approach for extracting the image's characteristics. The GLCM is a matrix with the same square matrix as gray levels in an image. These matrices built a 450-degree configuration. Contrasting, Homogenization, Chaos, Angle 2nd Momentum, and Similarity seem to be the statistical characteristics retrieved [22].

**Image categorization**. The process of categorizing images is known as classification. It divides oral images into malignant and non-cancerous categories. The malignant

**Fig. 6** LBP extracted oral image

and non-cancerous zones are classified using the DDM-FTIL classifier. The retrieved characteristics build a classification model for both malignant and non-cancerous areas. The predictor comprises five levels, each of which has 12 cells—the sample images as inputs to the algorithm. Images are categorized in Fig. 7. The classifier's return outputs 0 or 1. The healthy image is represented by pixel 0, whereas the aberrant malignant image is represented by pixel 1. The healthy image is represented by pixel 0, whereas the aberrant malignant image is represented by pixel 1.

**Image Segment Analysis**. The anomalous images must divide once they have been classified. The procedure of separating a malignant zone from a non-cancerous zone is known as fragmentation. First, we use morphology analysis to fragment the aberrant oral image. Next, the aberrant image is subjected to a dilatation and degradation procedure using a segmentation technique. Finally, a segmentation image is created by segmenting the damaged image from the expanded image. It has the benefit of being less time-consuming than segmentation as in Fig. 8.

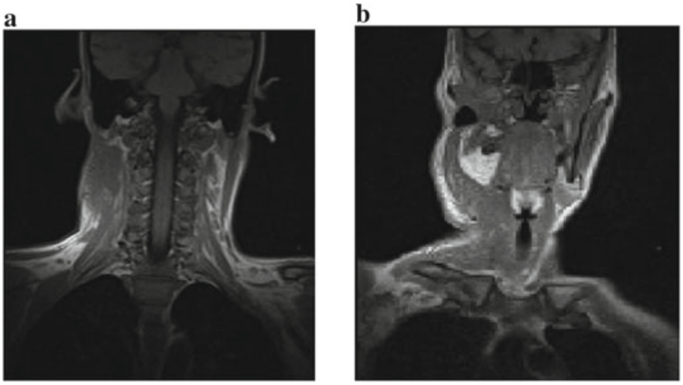


**Fig. 7** **a** Non-cancerous image, **b** cancerous images

**Fig. 8** **a** Region of interest, **b** morphologically treated image, **c** cancer zone spitted image

## 4 Experimental Results

The carcinogenic and non-cancerous portions of the oral cancer photos are preprocessed, categorized, and segmented. Then, the specificity, sensitivity, and accuracy metrics are used to assess the performance indicator (Table 1).

It shows the performance evaluation using the DDM-FTIL categorization technique gains 91.069% of sensitivity value, 91.844% value of specificity, and 94.386% of mouth cancer prediction accuracy as in Fig. 9.

Table 2 shows the results of Konstantinos et al. assessments [23], Anuradha et al. [4], Ahmed et al. [24], suggested classification. Compared to previous approaches, the suggested mouth cancer segmentation method obtains excellent performance

**Table 1** Performance assessment of mouth cancer categorization

| Image sequence | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---|---|---|---|
| 1 | 90.75 | 88.52 | 91.16 |
| 2 | 91.26 | 93.59 | 92.15 |
| 3 | 91.96 | 95.21 | 92.42 |
| 4 | 90.81 | 90.32 | 93.59 |
| 5 | 91.82 | 91.45 | 95.16 |
| 6 | 92.04 | 93.59 | 95.07 |
| 7 | 89.59 | 92.15 | 96.62 |
| 8 | 89.67 | 91.28 | 96.81 |
| 9 | 90.37 | 90.18 | 96.19 |
| 10 | 92.42 | 92.15 | 94.69 |
| **Avg** | **91.069** | **91.844** | **94.386** |

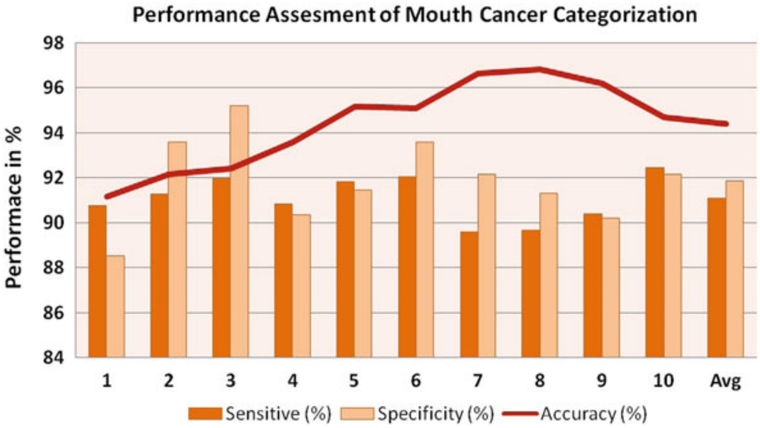


**Fig. 9** Performance evaluation of DDM-FTIL

**Table 2** Comparison of segmentation algorithms

| Methodologies | Threshold-based technique | Watershed segmentation algorithm | Decision support system | DDM-FTIL classifier |
|---|---|---|---|---|
| Sensitivity (%) | 87.45 | 90.17 | 90.19 | 92.14 |
| Specificity (%) | 87.75 | 91.76 | 88.45 | 93.71 |
| Accuracy (%) | 91.56 | 92.82 | 91.37 | 94.25 |

in various conditions of performance parameters such as sensitive, specific, and accurate, as shown in the table.

## 5 Conclusion and Future Enhancement

The research suggests utilizing a Gabor transform-based DDM-FTIL classifier to identify and segment oral cancer images. In the photographs, noise is eliminated and increased using the histogram equalization approach. Then, the improved image transforms into a Gabor image, and features (LBP and GLCM) are extracted. DDM-FTIL classifier uses to learn and classify these characteristics. The DDM-FTIL classification technique yields 93.71% sensitivity, 92.14% specificity, and 94.25% accuracy in segmentation. Our goal is to improve our technology to forecast many diseases in the future.

## References

1. Kaggle Inc. (2020) Oral cancer (lips and tongue) images. https://www.kaggle.com/shivam17299/oral-cancer-lips-and-tongue-images
2. Youssef B, Hamid Z, Abdelkhalak B (2016) Computer-aided diagnostics of facial and oral cancer. In: Third world conference on complex systems (WCCS). IEEE, Marrakech, Morocco
3. Chan C, Huang T, Chen C, Lee C, Chan M, Chung P (2019) Texture-map-based branch-collaborative network for oral cancer detection. IEEE Trans Biomed Circuits Syst 13(4):766–780
4. Anuradha K, Sankaranarayanan K (2012) Detection of oral tumor based on marker—con-trolled watershed algorithm. Int J Comput Appl 52(2)
5. Ghosh S, Biswas S, Sarkar D, Sarkar P (2014) A novel neuro-fuzzy classification technique for data mining. Egypt Inf J 15(3):129–147
6. Jung W, Zhang J, Chung J (2005) Advances in oral cancer detection using optical coherence tomography. IEEE J Sel Top Quantum Electron 11(4):811–817
7. Ram S, Siar CH (2005) Chemiluminescence as a diagnostic aid in the detection of oral cancer and potentially malignant epithelial lesions. Int J Oral Maxillofac Surg 34(5):521–527
8. Zare MR, Mueen A, Seng WC (2013) Automatic classification of medical X-ray images using a bag of visual words. IET Comput Vis 7(2):105–114

9. Xu Y, Mo T, Feng Q, Zhong P, Lai M, Chang EC (2014) Deep learning of feature representation with multiple instance learning for medical images analysis. In: 2014 IEEE international conference on acoustics, speech and signal processing, pp 1626–1630
10. Wan J, Wang D (2014) Deep learning for content-based image retrieval: a comprehensive study. In: Proceedings of the 22nd ACM international conference on multimedia, pp 157–166
11. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
12. Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M (2014) Medical image classification with convolutional neural network. In: Proceedings of 13th international conference on control automation robotics & vision (ICARCV). IEEE, Singapore, pp 844–848
13. Zeng R, Wu J, Shao Z, Chen Y, Chen B, Senhadji L, Shu H (2016) Color image classification via quaternion principal component analysis network. Neurocomputing 216:416–428
14. Chan T, Jia K, Gao S, Lu J, Zeng Z, Ma Y (2015) PCANet: a simple deep learning baseline for image classification? IEEE Trans Image Process 24(12):5017–5032
15. Bar Y, Diamant I, Wolf L, Greenspan H (2015) Deep learning with non-medical training used for chest pathology identification. In: Hadjiiski LM, Tourassi GD (eds) Proceedings of SPIE medical imaging, vol 9414. International Society for Optics and Photonics, Orlando, FL, USA, p 94140-V1-V7
16. Shin H, Roth H, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection-CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35(5):1285–1298
17. Rakotomamonjy A, Petitjean C, Salaun M, Thiberville L (2014) Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images. Artif Intell Med 61(2):105–118
18. Bruna J, Mallat S (2013) Invariant scattering convolution networks. IEEE Trans Pattern Anal Mach Intell 35(8):1872–1886
19. Cruzroa A, Basavanhally A, Gonzalez F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A (2014) Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Gurcan MN, Madabhus A (eds) Proceedings of SPIE, vol 9041, Houston, TX, USA, p 904103-1-15
20. Raja JV, Khan M, Ramachandra VK, Al-Kadi O (2012) Texture analysis of CT images in the characterization of oral cancers involving buccal mucosa. Dentomaxillofacial Radiol 41(6):475–480
21. Cen S, Yu Y, Yan G, Yu M, Yang Q (2020) Sparse spatiotemporal descriptor for micro-expression recognition using enhanced local cube binary pattern. Sensors 20(4437)
22. Baraldi A, Parmiggiani F (1995) An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. IEEE Trans Geosci Remote Sens 33:293–304
23. Exarchos KP, Goletsis Y, Fotiadis DI (2012) Multiparametric decision support system for the prediction of oral cancer reoccurrence. IEEE Trans Inf Technol Biomed 16(6):1127–1134
24. Ahmed CRM, Madeshan N, Kalaivanan S, Narayanan KS (2017) To detect and classify oral cancer in MRI image using firefly algorithm and expectation maximization algorithm. Int J Pure Appl Math 116(21):149–154

# Challenges in Digital Health Care—A Hybrid-Decision Approach Understanding Key Challenges in India

**Sachi Nandan Mohanty, Tilottama Singh, Richa Goel, and Sukanta Kumar Baral**

**Abstract** The pandemic has formed an urgent demand for coordinated systems to respond to the COVID across health sectors, and digital health resolutions have been recognized as promising methods to tackle this challenge. This paper discusses the current challenges in way of implementing a sound digital health ecosystem ensuring sustainable solutions in health networks. However, the appropriateness of digital emulsions may face challenges. More intensive measures must be implemented to make sure that future digital health programs will have a bigger impact on the pandemic and meet the most tactical needs to ease the life expectancy of people at the vanguard of the crisis. The key challenges on the road to implementing sound digital health in the Indian health care system have been found through a vast literature review and participants' opinions. The current study adopts a mixed methodology where interviews of 25 health care top management leaders are carried out for triangulation purposes. Understanding the key challenges of strategic hospitals in Pare is taken from the study. This leads to an exploration of the critical challenges on the road to the successful adoption of digital health. The findings are categorized into various themes through thematic analysis. The emergent themes provide the critical challenges in the successful implementation of digital health care in India. Further, it suggests an integrated framework by using the DEMATEL technique. The study's findings can be used by researchers, health care organizations, practitioners, and stakeholders working in the digital health arena for the betterment of society.

---

S. N. Mohanty
Vardhaman Engineering College, Hyderabad, India

T. Singh
UIM, Uttaranchal University, Dehradun, India
e-mail: tsingh6@amity.edu

R. Goel (✉)
Amity International Business School, Amity University, Noida, India
e-mail: rgoel@amity.edu

S. K. Baral
Department of Commerce, Faculty of Commerce and Management, Indira Gandhi National Tribal University (Central University), Amarkantak, Madhya Pradesh, India

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_9

**Keywords** Digital · Health · Service · Challenges

# 1 Introduction

Digitalization of health care has been the most promising and rewarding arena concerning attaining sustainability in health care operations. The road toward health care digitalization has glaring impacts on improving the health services but also faces tremendous challenges when it comes to ground implementation [1]. Little research has focused on digital health care challenges in emerging economies like India. Identifying the challenges experienced by health care with digitalization will make it possible for health care workers to work more effectively providing the services to the people. Also, the study provides the grounds for effective implementation of digital health care by addressing the key challenges in the health networks. It also fills the research gap by identifying the challenges to successful implementation of digital health care in India. The study findings will help in preparing a more conducive digital health care setup in India. Therefore, the present study highlights the following questions.

## *1.1 Research Questions*

- What are the major challenges restricting the effective implementation of digital health care in India?
- What are the most influential challenges in digital health care in India?
- What is the cause-effect relationship between the challenges based on expert opinion and theoretical underpinning?

Based on the research questions, the following objectives have been formulated.

# 2 Research Objectives

- To identify and scrutinize the major challenges in digital health care with context to India
- To find the most influential challenges in digital health care.
- To state the causal and effective challenges in digital health care considering expert opinion and theoretical underpinning?

The digitalization of health care during and post the COVID-19 has seen many challenges. The future predictions in the digital health care arena symbolize potential growth in future areas with massive adoption of digital health gadgets as depicted in the figure below (see Fig. 1).

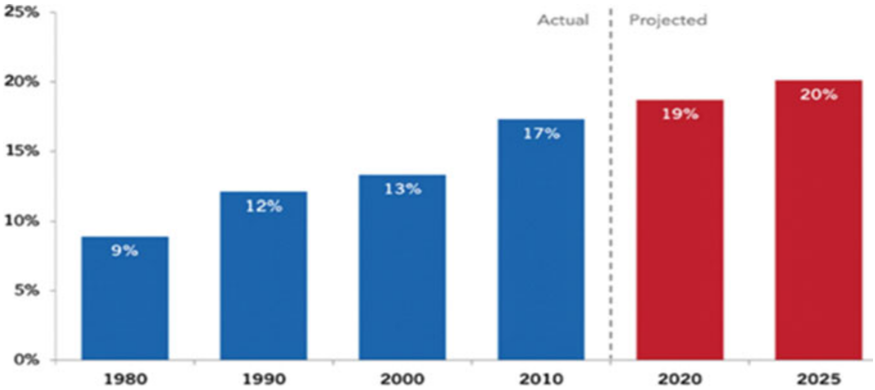


**Fig. 1** Peterson, 2019

However, the predominance of challenges is an area of exploration. There are specific to the Indian context that if not addressed effectively might turn out to be huge hurdles in reaping effective benefits of digital health care in India (Figs. 2 and 3).

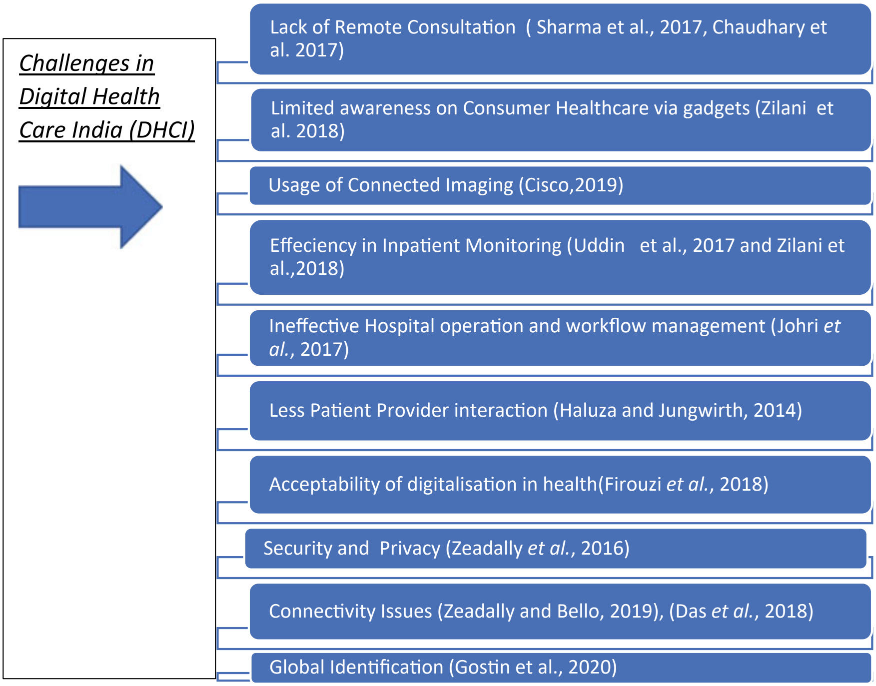


**Fig. 2** Projected growth in digital health care

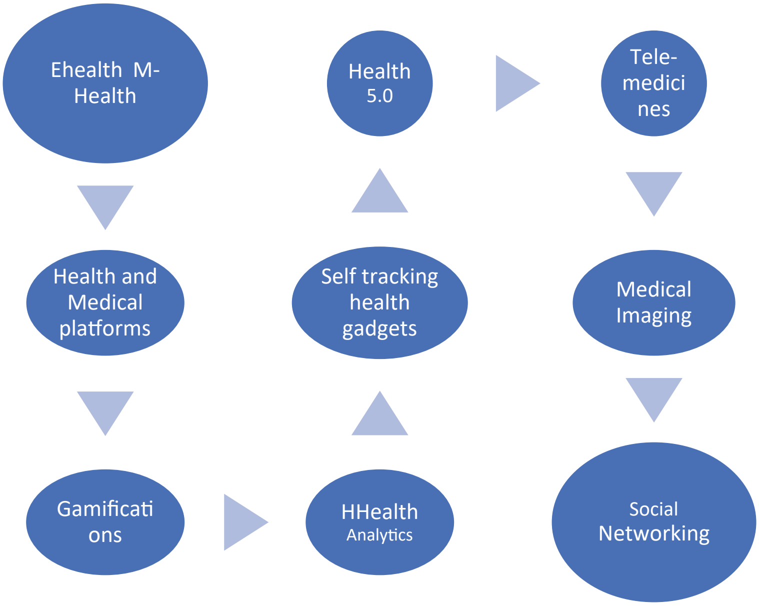


**Fig. 3** Challenges in digital health care. *Source* Based on Literature

The COVID-19 digital backup relies upon medical care, financial, and social viewpoints, proficiency, viability, and medical services conveyance dependability [2]. Studies have contemplated a choice of an emotionally supportive network for dealing with the interest in the medical care store network to moderate the medical services production network interruptions during a pandemic. Another study by Zhu et al. [3] proposed bled health care network digital heath care though has been an instigator in bringing several gadgets on the ground making life easier but still has a long way ensuring implementation.

Thus, the major challenges extracted are referred to as Critical Challenges (CR) like CR1 Lack of remote Consultation, CR2 Limited awareness of consumer health care via gadgets, and so on. Subsequently, these challenges were scrutinized using the thematic analysis, which helped estimate the most pertinent challenges based on the interviews and survey conducted. The participants are chosen from the medical industry as purposive sampling.

## 3 Research Methodology

For the study, interviews and surveys were conducted to delve into the issues concerning the challenges in digital health care effectiveness. The interviews and survey were the most suitable choice for the study as they witnessed the actual

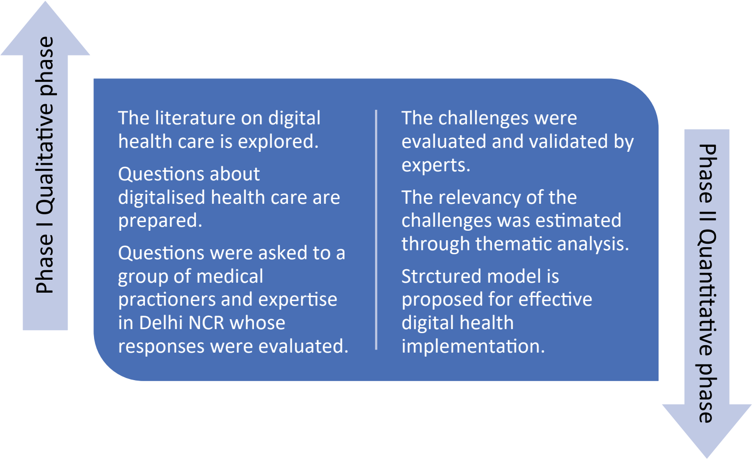


**Fig. 4** Research methodology framework

hurdles faced by the relevant group in the digital health care platforms. The interview sessions were conducted with health care providers working in Delhi-NCR. All the responses were analyzed using thematic analysis. The coding was developed and organized to assimilate the emerging themes from the interviews. The findings were further presented in a tabular form for necessary interpretation (Fig. 4).

## 4 Result Analysis and Discussion

The experts have been asked to rate all the criteria for evaluating a digital health care system, and their ratings have been recorded (Fig. 5 and Table 1).

The majority of the respondents 80.2% have stated that the need for global health care identification, which can connect the residents globally with an ID is a must to reap extensive benefits of the digital health care platform. In addition, 73.2% of respondents have shown that hospital operation and workflow management act as a potential bottleneck in implementing a strong digital interface with the patients. The third major area coins as the acceptability of the digital frames by people across rural and urban areas. The major challenge seen is in rural areas when it comes to the adoption of such digital apps, checkups, and aid. This needs special attention for broader acceptability and adoption. The three major challenges further followed by connectivity (71.2%) issue, privacy challenges (63.9%), inpatient monitoring (62.6%), remote consultation (61.2%) then telemedicine, and consumer health care where the least weightage was given to connected imaging (42.2%), being least significant.

Thus, the findings reveal that the global identification ID is a major challenge followed by effective operation and workflow management. The result of the study

**Fig. 5** Word cloud

**Table 1** Challenges in digital health care in India

| Challenges in digital health care | Percentage |
|---|---|
| Global identification | 80.2 |
| Connectivity issue | 71.2 |
| Privacy and security | 63.9 |
| Acceptability of digital health care | 73 |
| Telemedicine | 56.7 |
| Remote consultation | 61.2 |
| Consumer health care | 54.4 |
| Connected imaging | 43.2 |
| Inpatient monitoring | 62.6 |
| Hospital operations and workflow management | 73.2 |

can encourage the hospital management teams and respective role players to significantly work on the challenges making the implementation of digital health care stronger.

Step 1: Identification of the evaluation criteria
Step 2: Determining direct relation matrix

$$A^k = \begin{array}{c} C_1 \\ C_2 \\ \vdots \\ C_n \end{array} \begin{bmatrix} 0 & \otimes a_{12}^k & \dots & \otimes a_{1n}^k \\ \otimes a_{21}^k & 0 & \dots & \otimes a_{2n}^k \\ \vdots & \vdots & \dots & \vdots \\ \otimes a_{n1}^k & \otimes a_{n2}^k & \cdots & 0 \end{bmatrix}$$

Step 3: Combination of all DR Matrices
Step 4: Determining, the structural model

$$\sum_{j=1}^{n} \otimes z_{ij} = \left[ \sum_{j=1}^{n} z_{ij} \ , \sum_{j=1}^{n} \overline{z}_{ij} \right]$$

and $r = \max_{i<j<n} \left( \sum_{j=1}^{n} \overline{z}_{ij} \right)$

$$G = \begin{bmatrix} 0 & \otimes g_{12} & \dots & \otimes g_{1n} \\ \otimes g_{21} & 0 & \dots & \otimes g_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \otimes g_{n1} & \otimes g_{n2} & \cdots & 0 \end{bmatrix}$$

$$\otimes g_{ij} = \frac{\otimes z_{ij}}{r} = \left[ \frac{z_{ij}}{r} \ , \frac{\overline{z}_{ij}}{r} \right]$$

Step 5: To establish the Total relation matrix:
Step 6: Whitenization
Step 7: Result analysis: Assimilating the values of $d$ and $r$ (Table 2).

The data was collected from top management leaders of prime hospitals on the top management level AIIMS, Ram Manohar Lohia hospital, Sardar Vallabh Bhai Patel Hospital, Safdarjung hospital, Rajiv Gandhi, Ganga Ram, and Delhi Trauma Center. The sample was taken from premiere strategic hospitals. To examine the DEMATEL further, where a sample of 25 experts was taken (Table 3).

Further to explore the cause-effect relationship amongst the challenges, the result was obtained by calculating the $D + R$ and $D - R$ where the positive values denote the cause and negative values denote the effect.

## 5 Conclusion

The relationship is established by getting out the line value and then evaluating the values which are > than the threshold value in each critical challenge factor listed. As shown in Table 4 we can state that CR1, CR2, CR8, and CR10 are effect challenges while CR3, CR4, CR5, CR6, and CR9 are causes. This helps us understand the

**Table 2** Values of $D$ and $R$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | [1.000, 1.000] | [1.136, 2.024] | [1.016, 2.014] | [1.343, 1.019] | [1.106, 3.104] | [1.116, 0.004] | [0.116, 2.004] | [1.216, 1.014] | [0.101, 0.125] | [1.106, 1.014] |
| 2 | [0.016, 0.014] | [0.000, 0.000] | [0.133, 0.159] | [0.189, 0.010] | [0.216, 0.104] | [0.216, 0.140] | [0.103, 0.018] | [0.042, 0.114] | [0.143, 0.039] | [0.143, 0.069] |
| 3 | [0.000, 0.000] | [0.122, 0.025] | [0.000, 0.000] | [0.326, 0.114] | [0.143, 0.059] | [0.033, 0.059] | [0.033, 0.039] | [0.016, 0.124] | [0.123, 0.218] | [0.147, 0.049] |
| 4 | [0.143, 0.029] | [0.123, 0.024] | [0.126, 0.124] | [0.000, 0.000] | [0.216, 0.114] | [0.133, 0.135] | [0.133, 0.019] | [0.257, 0.115] | [0.012, 0.035] | [0.033, 0.065] |
| 5 | [0.011, 0.025] | [0.136, 0.051] | [0.100, 0.015] | [0.236, 0.114] | [0.000, 0.000] | [0.113, 0.135] | [0.327, 0.129] | [0.124, 0.235] | [0.156, 0.065] | [0.000, 0.000] |
| 6 | [0.100, 0.025] | [0.123, 0.065] | [0.243, 0.059] | [0.134, 0.035] | [0.016, 0.104] | [0.000, 0.000] | [0.114, 0.035] | [0.226, 0.104] | [0.067, 0.073] | [0.171, 0.030] |
| 7 | [0.000, 0.016] | [0.226, 0.104] | [0.043, 0.047] | [0.136, 0.104] | [0.126, 0.104] | [0.116, 0.114] | [0.000, 0.000] | [0.236, 0.104] | [0.343, 0.099] | [0.456, 0.440] |
| 8 | [0.000, 0.000] | [0.123, 0.349] | [0.157, 0.129] | [0.343, 0.039] | [0.100, 0.045] | [0.000, 0.034] | [0.211, 0.100] | [0.000, 0.000] | [0.010, 0.025] | [0.046, 0.000] |
| 9 | [0.221, 0.000] | [0.133, 0.049] | [0.226, 0.114] | [0.133, 0.049] | [0.116, 0.065] | [0.143, 0.059] | [0.111, 0.135] | [0.156, 0.104] | [0.000, 0.000] | [0.000, 0.004] |
| 10 | [0.239, 0.129] | [0.179, 0.129] | [0.136, 0.114] | [0.113, 0.108] | [0.133, 0.039] | [0.129, 0.126] | [0.033, 0.089] | [0.223, 0.148] | [0.178, 0.104] | [0.000, 0.003] |

**Table 3** Average matrix

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2.65 | 2.75 | 2.6 | 2 | 2.8 | 2.6 | 2.8 | 2 | 1.7 |
| 2.3 | 0 | 2.6 | 2.65 | 2.7 | 2.6 | 2.3 | 2.7 | 1.6 | 2 |
| 2.35 | 3.15 | 0 | 2.75 | 3.15 | 3.35 | 3 | 3.2 | 3 | 2 |
| 2.8 | 2.7 | 3.2 | 0 | 3.1 | 2.75 | 1.5 | 1.8 | 2 | 3 |
| 2.35 | 1.7 | 2.25 | 2.35 | 0 | 2.7 | 2.9 | 2.6 | 2.8 | 2.9 |
| 2.85 | 2.15 | 1.55 | 1.8 | 1.75 | 0 | 3 | 2.5 | 2.8 | 3 |
| 2.67 | 2.14 | 1.65 | 1.7 | 1.65 | 1.7 | 0 | 3 | 2.7 | 3 |
| 2.76 | 2.15 | 2.5 | 2.6 | 2.7 | 1.8 | 2 | 0 | 2 | 3.1 |
| 2.8 | 27 | 2.3 | 26 | 25 | 2.6 | 2 | 3 | 0 | 3.2 |
| 2.6 | 2.5 | 2.8 | 2.9 | 3 | 2 | 3 | 2.1 | 2.7 | 0 |

**Table 4** Cause and effect relation results

| Challenges | $D$ | $R$ | $D + R$ | $D - R$ |
|---|---|---|---|---|
| CR1 | 4.02 | 7.23 | 11.25 | −3.21 |
| CR2 | 5.25 | 6.01 | 11.26 | −0.76 |
| CR3 | 8.03 | 5.04 | 13.07 | 2.99 |
| CR4 | 8 | 5.04 | 13.04 | 2.96 |
| CR5 | 6.78 | 6.3 | 13.68 | 0.38 |
| CR6 | 7.07 | 6.61 | 13.68 | 0.46 |
| CR7 | 6.07 | 5.12 | 11.19 | 0.95 |
| CR8 | 7.2 | 8.2 | 15.4 | −1 |
| CR9 | 8.1 | 7 | 15.1 | 1.1 |
| CR10 | 7.23 | 8 | 15.23 | −0.77 |

relationship between the given factors more in detail highlighting the causal and effective factors. Also, the thematic analysis has showcased these effect challenges to have a major dominance being the resultant of the causal challenges.

## 6 Future Recommendation

By evaluating the huge amounts of recorded patient data created by high-tech devices from various sources, alphanumeric health will assist the future demands of medicine [4]. They also require comprehensive and effective human validation, as well as increased clinical dependability. Medical practitioners must also become aware of and adapt to these advancements to provide better health care to patients. Researchers, scientists, physicians, payers, and regulators must work alongside technology developers to achieve the ultimate objective of assisting patients in living longer and feeling better. The study aims to suggest an integrated framework aiming to lessen the challenges in the successful implementation of digital health care. The proposed framework is given below which rests on major stakeholders influencing the efficiency of digital health care which are the government, international and national health agencies, health aid manufacturers, and the research and development, team (Fig. 6).

The study aims to provide critical insight into the relevancy of challenges in digital health care in India and to propose an integrated framework showcasing the impertinent steps for the effective implantation of a digital health setup in such unprecedented times.

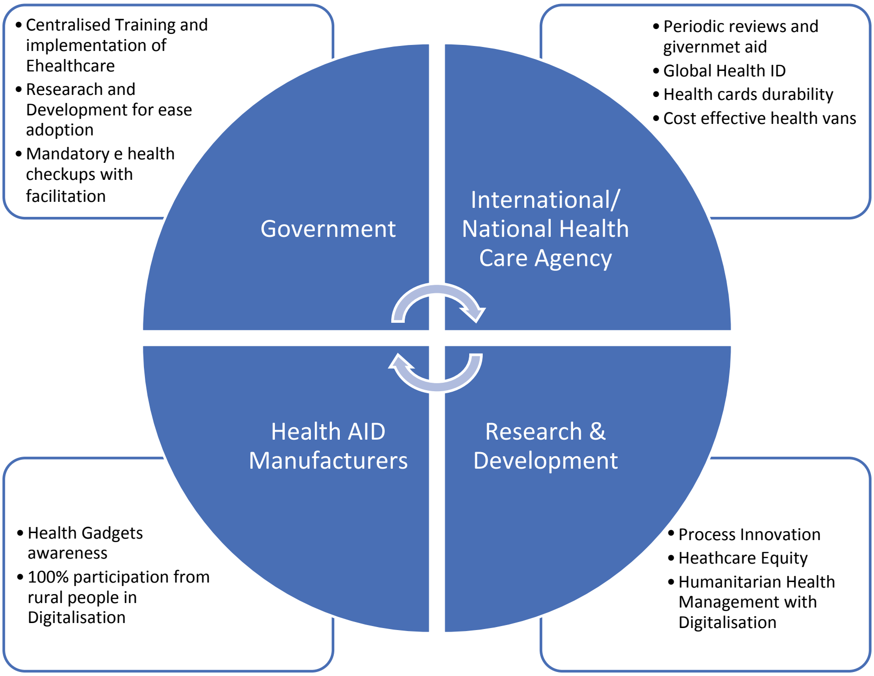


**Fig. 6** Integrated digital health care network. *Source*: Author's own

## References

1. Praveena MA, Bharathi BB (2017) A survey paper on massive information analytics. In: International conference on information communication and embedded systems (ICICES). IEEE, pp 1–9
2. Haleem A, Javaid M, Vaishya R (2020) Effects of Covid-19 pandemic in day-by-day life. Curr Remedy Stud Pract 10(2):78–80
3. Zhu L, She ZG, Cheng X, Qin JJ, Zhang XJ, Cai J, Li H (2020) Association of blood glucose manage and results in sufferers with Covid 19 and present kind II diabetes. Cell Metab 31(6):1068–1077
4. Abbas S, Hadi AA, Abdullah HO, Alnoor A, Khattak ZZ, Khaw KW (2021) Encountering Covid 19 and perceived pressure and the function of fitness weather amongst scientific workers. Curr Psychol 1:1–14

# Telemedicine and Healthcare Setting for Remote Care of Patients: Status and Future Implications

**Sachi Nandan Mohanty, Venaik Anita, Sukanta Kumar Baral, Richa Goel, and Tilottama Singh**

**Abstract** One of the best technologies is Telemedicine (TM) which has recently gained a lot of popularity because of its intensifying deployment and development of digital amenities, worldwide. TM is associated with various healthcare services like medicine, surgery, diagnosis, metabolic, etc. and provide access to patients' care. Also enhances resource and reduced treatment costs. Since TM have good impacts on the better clinical outcomes, better patient happiness, better accessibility and healthcare systems. Even, during current pandemic (COVID-19) where public health emergency happened, then TM had act as a viable option. TM acted as safest shield between patients (either infected or uninfected) and healthcare experts for interactive health consultant. So that, its maintained social distance to prevent the spreading of this communicable clinical issue. Therefore, the current study tried to describe about the role of TM in the healthcare area and its efforts during the current pandemic.

**Keywords** Telemedicine · Digital · Healthcare experts · Covid-19 · Pandemic

---

S. N. Mohanty
Vardhaman Engineering College, Hyderabad, India
e-mail: sachinandan09@gmail.com

V. Anita · R. Goel
Amity Business School, Amity University, Noida, India
e-mail: avenaik@amity.edu

R. Goel
e-mail: rgoel@amity.edu

S. K. Baral (✉)
Department of Commerce, Faculty of Commerce and Management, Indra Gandhi National Tribal University (Central University), Amarkantak, Madhya Pradesh, India
e-mail: sukanta.baral@igntu.ac.in

T. Singh
UIM Uttaranchal University, Dehradun, India
e-mail: tilottamasingh2101@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_10

## 1 Introduction

Telemedicine (TM) is a digital system that uses a combination of electronic devices and digital information. It is based on the inaccessible communication mechanism. Although, TM entirely support healthcare services through "Remote Mode" [1]. It is known by various terms like telehealth/TM consultation center (TCC)/TM specialty center (TSC)/TM system. This mode allows obtaining accurate information about patients and their health issues. However, it also plays a good role in a village or rural and remote areas worldwide that fight to get reliable, high-quality specialty medical care.

Globally, several medical institutes have backed TM technology and are reaping the benefits during the exposure time. Because TM has both a screening and diagnostic method that has gained significant grip; to enhance the distribution and increase the progress of Digi-health.

Since TM's inception during 1950 it has expanded its method in different departments of healthcare systems. The automated arsenal accessible to clinical fields has increased exponentially due to the progress and extensive social acceptance of this communicable technology like "Personal Computer-(PC)" and "Smart-Phone-(SP)".

"Smart-Phone-(SP)" and several healthcare providers have found that TM services are cost-effective and help physicians deliver better treatment. It has also improved resource clinical utilization and the low costs associated with conventional clinic visits. It has been used in various medical fields. Therefore, for all health interventions, skilled nursing facilities, including "Acute Care Hospitals", more than half of all acute care hospitals use TM incapacity. The TM system is expected to be worth $30 billion by 2019, rising at a rate of 20–50% per year and its practical clinical application expanding day by day. In general, TM can be used for primary care (non-emergency health issues or those that do not require the patient to be seen) of patients through phone calls or SPs. Still, it has not been shown that it can replace face-to-face consultations or outpatient departments (OPD) when needed but rather complements them [2]. A TM system may provide an immediate call to healthcare facilities when clinicians have closed their clinics. Consequently, technological evolution in the clinical setting is always focused on providing standard clinical care facilities (pre-post or without any urgent) to patients and this care with a particular prominence to patient and healthcare providers' pleasure, meanwhile maintaining optimal access to care; as a result, all health experts can continue to accept and use TM because of its many advantages.

***[This present paper has tried to describe the role of TM to emphasizing disease control during this pandemic.]***

***"History of Telemedicine"***

Telemedicine was a modern digital device technology used in some way before 30 years ago, but now (before 5–6 years ago), TM technology has established itself in a variety of scientific and biomedical fields, covering almost all areas. During flights, functional parameters (like physiological) were transmitted from both the spacecraft and the spacesuit. Even though data was collected and reported during this period, it was not connected to crew members in real-time during these early missions.

The projects have aim to facilitate healthcare services to the peoples of 5isolated/remote Gulf cities of the “Carpentaria Region (CR)”. In 1989, NASA was launched first foreign TM amenity, known as The Space Bridge to Armenia/Ufa. TM consultations or TM experts had performed their medical facilities with help of one-way camera during this project, speech, and likeness technologies between a healthcare setting center in Yerevan, Armenia, and 4 medical centers in the United States.

## *1.1 TM Adoption*

NASA had adoption, implementation of TM technology and then to begin as a revolutionary digital inkling. It has quickly evolved and gained acceptance by both patients and healthcare experts. The department of Veterans Affairs of the “United States (US)” had provided excellent Telemedicine related healthcare services from the past two decades. In 2014, 44 different healthcare experts had provided 1.2 million TM healthcare facilities to virtually 700,000 Veterans. In addition, the “Veterans Affairs (VA)” system has made TM a priority in providing better treatment to its veterans. Currently, TM is an important healthcare tools with high success target for fiscal years 2014–2020 and further, in the 2016 for TM functions, approximately, one point two billion dollar set aside “VA” budget.

## *1.2 Status of TM in India*

Every person in Utopia should have immediate access to the right specialist for clinician advice; however, this cannot be a fantasy in the real world. All people are equal, but some are more equal than others, as the saying goes; however, in India, still a day there is not possible rout to deliver complete “Primary Medical Care (PMC)” services in villages as like same, in sub-urban and urban areas.

Also, incentives to entice professionals to practice have failed. Although India’s computer literacy is rapidly increasing and will talk about TM is now being seen as a new “Avatar” for healthcare specialists. Un-doubt, theoretically, this is much easier to build a model for telecommunications system or TM technology in peripheral and rustic region of India than it is to send one hundred medical experts there; whether, at the same time, “Satellite-Based Technologies (SBT)” and “Fiber Optic Cables” are the way of the future in telecommunications [3].

The Indian healthcare system has modified because of various developed advanced TM software and the first time—Center for Development of Advanced Computing-India (C-DAC, India) has also lunched new TM implements or software services using public and personal cloud computing services. Further these softerwares are opinionated to the different departments of medical organization or hospitals or clinical associations such as Tele-Cardiology (TC), Tele-Radiology (TR), and Tele-Pathology (TP) etc.

"Apollo group of hospitals, Andhra Pradesh" was initiating a pilot project at a secondary level hospital in Aragonda as a Trailblazer a village 16 km from Chittoor and their population 5000, Aragonda project, as a result the local healthcare experts of hospital are using various digital communication tools or software like video conferencing technical tools for the treatment of patients at concerned stages.

In addition, "Indian Space Research Organization (ISRO)", had established Very Small Aperture Terminal (AVSAT) satellite which began with simple webcams and ISDN telephone lines. Further, "Sriharikota Space Center project" located 130 km away from Chennai, served as an important Launchpad for the Indian Space Research Organization in this field. According to recent survey, TM association or network of ISRO has expanded and connected with 45 remote and rural hospitals and also associated with fifteen super-specialty medical organization. However, TM is also very useful to other region (hard to reach area Islands of Andaman and Nicobar and Lakshadweep, mountainous and hilly districts of Jammu and Kashmir, including Kargil and Leh, Medical College hospitals in Orissa, and several rural/district hospitals in mainland states) of India, are among the remote or rural nodes. Currently, the Indian government has actively supported various TM programs by the mentioned organization, shown in Fig. 1.

Although in India, a governmental or private organization actively sponsors (Fig. 1) TM services in for patients' care. Also, India wants to share TM expertise or services with other country like African countries still where PHC infrastructure is not started of basic healthcare is not well setup. For this same purpose, India has launched and successfully implanted "Pan African e-network Project" that connecting fifty-three African hospitals to eleven specialized tertiary healthcare hospitals (TCH).

After the successful implementation and positive end come—TM network was multiplied and connected with Southeast Asia SAARC's countries. Also, through the e-Network of the South Asian Association for Regional Co-operation (SAARC).

**Fig. 1** Various TM programs supported by government of India

Government of India has initiated and running TM project to improve access healthcare facilities. All SAARC nations will benefit from specialized healthcare facilities and treatment. It will also aid healthcare professional “Bhutan, Nepal, and Afghanistan” continuing medical education in SAARC countries and already they have signed a memorandum of understanding, and several projects have begun.

## 1.3 Modalities of Telemedicine

The healthcare setting revolution has changed the structure of health organizations for better treatment of patients through TM technology. This acts as a stable and secure broadband link essential for providing effective long-distance healthcare facilities to patients. A robust broadband signal transmission system, also recognized as a “line”, consists of complex infrastructure that allows data to flow at low speeds. Globally, recent advancements in these healthcare substructure facilities have enabled cost-effective access to wider bandwidths, allowing for faster documents transfer speeds. In the current situation, TM technology has been executed for healthcare services through various modalities like “Short message service—SMS” via typescript messaging (this is an essential modality that has been used for a long time), SP applications, automated calls, and digital wearable gadget.

Although it was the fact that already SMS message has been actively used to mobile health (*M*-health-Fig. 2) apps and fitness trackers which act as a fastest-growing segments driving TM modality. Now a day these different TM modalities or applications can provide potential medical care services through sharing of pictures and also a video with audio. Despite its benefits, some physicians argue that app-based modalities may create a barrier to remote healthcare delivery because it is not necessary that all patients have access to “SP” and “PC” which are compulsory to use TM related digital map.

A few concerns arise regarding the underuse above mentioned digital tools and apps (mobile phone (MP) and *M*-Health apps) by certain populations (the elderly and those from poorer socioeconomic backgrounds) may be close, because overall

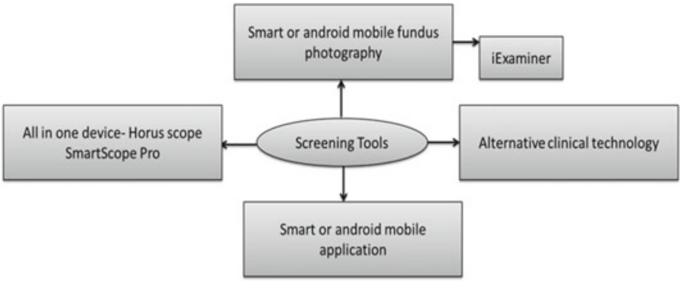


**Fig. 2** TM networks for the screening/identifying or the diagnosis of clinical disease or disorders

these tools or apps have emerged like viable options for reducing hospital stays. Also, improving the quality of patient care by earlier detecting patients' complications. As a result, reducing monthly re-consultancy, fast improving/or recovery, and lowering healthcare costs in the treatment of patients.

As TM becomes more con-joint matters such as how providers are compensated may need to be addressed and it can be a tough decision to make. Then there's the clinical risk of being too reliant on this TM scheme. Due to the risks associated with determining what information is accurate versus inaccurate, overdependence or overuse of TM will quickly spiral out of control unless more uniform policies and procedures are implemented.

TM technology has yet to develop to the point where it can be employed on a continuous or flexible basis. Although, it has the potential to be a huge benefit to the planet and all of its civilizations. The contributions it will make are limitless. However, TM would need more time and effort to plan before it can be confidently used. Different technologies' methods of execution—TM can provide different clinical care services with healthcare experts in the hospital through three main categories.

## *1.4 Current Application of TM*

### 1.4.1 Significance of TM During This Pandemic "COVID-19"

'COVID-19 is a global crisis affecting entire community and arises primary healthcare among the pandemic seem to be a biggest healthcare challenge, not only in India but also at worldwide. SARS-CoV2, is a highly infectious disease which caused "COVID-19" (most communicable disease), therefore World Health Organization (WHO) has declared this is pandemic.

Globally, SARS-CoV2 has infected more than four hundred million individuals; resulting high mortality rate (more than 175,000 deaths). The number of infected people and death rate are increased day by days but there is not possible treatment except safety measure like distancing oneself from others is thought to be effective or helping in reducing the rate of man-to-man transmission and diminishing morbidity and mortality rate worldwide. The capacity and accessibility of healthcare facilities are so inadequate in low-resource/or developing countries. A number of hospitals/healthcare settings are facing challenges during this pandemic because of lack healthcare tools like ventilators, ICU beds, and medical staff. Additionally, personal protective equipment (PPE) for healthcare workers is not enough. Healthcare workers are becoming infected due to a shortage of safety and medical tools. Many hospitals and clinics are hesitant to provide healthcare in this situation because they are afraid of being exposed to the coronavirus. As a result, even individuals with non-COVID illnesses (acute and chronic illnesses, such as diabetes, pregnancy, obesity, malnutrition, chronic respiratory diseases, cardiovascular disease, cancer, and mental health conditions) are regretted to access for healthcare services from hospitals or medical care centers and difficult to meet healthcare needs during this pandemic. Therefore, during this crisis or pandemic, TM and telehealth can play a critical role in ensuring PHC needs or services or healthcare experts consult to acute/chronic clinical issues or consult with patients from afar.

The services of TM technology are based on the 2 way of communication mode like information and communication technology (ICT) that exist between a patient and healthcare experts. Remotely, the smart healthcare experts communicate through video conferencing using Tablets, or Computers and also MPs and provide best the patient health advice and digital prescriptions. This allows the experts and the patient to see each other, giving the impression like "Real or Physical Visit" to medical organization.

TM's utility in the face of a pandemic is a topic that has been discussed in the context of addressing these issues during this pandemic. Restrictions on human mobility or lockdown can have a significant impact to stop the spread of new coronaviruses' second and third strains or waives. The World Health Organization (WHO) has urged all countries to take all necessary strict steps to prevent infection. Lockdown or mobility restrictions have become a barrier for patients suffering from a variety of acute and chronic illnesses in personally visiting healthcare organizations. Being unable to obtain medical assistance for an extended period can put their health at risk. Intervention of TM technology can aid patients avoid complications later on, so that, TM experts can be provide best care services to patients for maintaining a healthy lifestyle during. Due to a lack of shortage of testing kits (RT-PCR kit) in developing countries, investigation facilities are so restricted, and numerous positive patients go un-tested and they can also infect other people.

Non-infected patients are highly sensitive to coronavirus when they will visit medical center: although, other patients than COVID-19, can be protect by using TM technology. Furthermore, TM consultations could be more effective in terms of preventing infection spread and also with mild corona positive patients who does not require hospitalization and also these services can shrink the requirement of PPEs [4].

On the other hand, these home-confined health experts can make good use of their time by providing TM consultations to corona positive patients the majority of the population in many developing countries—(India, Bangladesh, and Pakistan, still lives in rural or villages. However, generally medical care experts work in urban areas, therefore, shortage of qualified clinician in villages, resulting in a geographic imbalance in healthcare experts' availability. Thus, TM services can decrease the burden and assist healthcare experts to evenly distribute clinical services across the nation. Now several current reports have declared that the COVID-19 infection is not equally affected in all countries around the world. It is not evenly distributed in all regions of the same country. So that in this case, TM can help healthcare experts manage their workload by allowing any healthcare experts to provide primary health consultations regardless of their location through various digital modules.

**Table 1** General challenges associate with the development of TM

| | |
|---|---|
| Healthcare experts' perspectives | Doctors are unsure about e-medicine and are unfamiliar with it |
| Financial unavailability | Telemedicine is often financially unfeasible due to high technology and connectivity costs |
| Patients' anxiety and unfamiliarity with e-Medicine | Patients' lack faith in the outcome of e-medicine |
| Lack of best necessities | Almost 40% of India's population lives in poverty. Transportation, power, telecommunications, clean drinking water, primary healthcare and other basic facilities are all lacking. When an individual has nothing to improve, no technological innovation can help |
| Lack of awareness and basic infrastructure | Lack of adequate telecom infrastructure may also be a big stumbling block in this regard, as telecommunication necessities better networks and higher internet speeds, which can be a huge challenge in and themselves |
| Regulations | While the Indian government recently released telemedicine practice guidelines, the regulations that will enable telemedicine to be widely used are still being produced. There was no regulation on the practice of telemedicine prior to the announcement of the new guidelines, which generated ambiguity for doctors and medical professionals while casting doubts on the entire procedure |

## 2 Challenges of Telemedicine Technology

Although digitalization has led to us preferring unconventional approaches that are more convenient, they often come with their own set of problems. Lack of adequate telecom infrastructure may also be a big stumbling block in this regard, as teleconsultation necessitates better networks and higher Internet speeds, which can be a huge challenge (Table 1) in and of themselves.

## 3 The Future Aspect of Telemedicine Technology

TM technology has the potential to be a compelling alternative to conventional medical care services and preventative health issues while also improving clinical outcomes. It is anticipated to continue to move healthcare service delivery from healthcare system or clinic into the residence in the developed countries.

It will mostly be employed in clinical applications that connect providers stationed in the health system-referral hospitals, and tertiary centers in the poor world or places with poor infrastructure. Then again it is also clear that the bright future of TM technology depends on human factors, economics, and technology [5]. However, this important point the future aspect of TM. Now a day, private organizational or societal changes are all influenced by TM technology. In the healthcare-related industrialized creation, personnel shortages, and declining 3rd-party re-imbursement are major drivers of technology-enabled healthcare, particularly in the fields of self-care. Various digital utensils-sensor devices including nanotechnology are all expected to alter the way of delivery clinical services [6]. In the coming generation, health system, patients, providers, and payers will benefit greatly from the rise and integration of information and communication technology into healthcare delivery.

In the economical view, the global TM market is expected to increase at a CAGR of 16.9% from USD 25.4 billion in 2020 to USD 55.6 billion by 2025, up from USD 25.4 billion in 2020. Some variables factors affect the TM market's growth, but one of the most important is the COVID-19 epidemic. 2020–2021 is undoubtedly one of the most eventful years in the history of telehealth. The future of TM has been revolutionized with COVID-19 and the rate of TM adoption has increased dramatically. It is proven that TM technology is playing a very important role in the healthcare setting. Over the last few years, the TM market has exploded. It allows healthcare practitioners to evaluate, diagnose, and treat patients virtually using telecommunications. According to Statista, North America is the world's largest TM market, with revenues expected to reach $35 billion by 2025. Also, according to a Mckinsey survey, the number of people who wish to use telehealth has increased dramatically from 11 to 76%.

## 4 Conclusion

In short, TM has grown to be one of the fastest-growing digital aspects of the healthcare setting or medical organization, globally [7]. Currently, TM play significant role in the different departments has gained prominence due to great clinical outcomes with high patient satisfaction. Also, reduced driving distance and wait times, and cost savings for both patients and healthcare experts.

The evolution in the medical care should continue to focus on offering the highest patient care quality with a high degree of patient satisfaction, yet allowing the greatest access to treatment. TM looks to achieve these objectives and should be viewed as a feasible option for all medical institute or organizations. According to recent clinical observation of survey, TM can play a significant role in this pandemic due to limiting viral spread, properly utilizing the time of healthcare experts, and treating health difficulties. Many sections of the developing countries are covered by mobile networks and internet app. All are done by help of TM. Future research will focus on patient confidentiality as well as the broader diffusion and use of TM in a healthcare setting.

## References

1. Rubulotta F, Soliman-Aboumarie H, Filbey K, Geldner G, Kuck K, Ganau M, Hemmerling TM (2020) Technologies to optimize the care of severe coronavirus disease 2019 patients for health care providers challenged by limited resources. Anaesth Analg
2. Sikkema RS, Pas SD, Nieuwenhuijse DF, O'Toole Á, Verweij J, van der Linden A, Koopmans MP (2020) COVID-19 in health-care workers in three hospitals in the south of the Netherlands: a cross-sectional study. Lancet Infect Dis 20(11):1273–1280
3. Chen PM, Hemmen TM (2020) Evolving healthcare delivery in neurology during the coronavirus disease 2019 (COVID-19) pandemic. Front Neurol 11:578
4. Govindan K, Mina H, Alavi B (2020) A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: a case study of coronavirus disease 2019 (COVID-19). Transp Res Part E: Logistics Transp Rev 138:101967
5. Vielma-Aguilera A, Bustos C, Saldivia S, Grandón P (2021) Psychometric properties of the attitudes scale of health care professionals' toward people with a diagnosis of mental illness (EAPS-TM). Curr Psychol 1–13
6. Roy D, Tripathy S, Kar SK, Sharma N, Verma SK, Kaushal V (2020) Study of knowledge, attitude, anxiety and perceived mental healthcare need in Indian population during COVID-19 pandemic. Asian J Psychiatry 51:102083
7. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S (2015) Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database of Syst Rev (9)

# Performance Evaluation of ML Models in the Health Care Domain to Classify Liver Disease: A Case Study

**Goutam Datta and Anupam Singh**

**Abstract** In the Artificial Intelligence (AI) field, we have seen a massive paradigm change. The transition from the earlier rule-based approach in expert systems to today's data-centric approach in machine learning (ML) and deep learning has drastically altered the implementation scenario of many AI-based applications. MYCIN was a rule-based expert system that we used in the past to detect bacteria causing disease. In today's AI, however, machine learning and deep learning-based technologies are frequently used in the health care industry. When there is enough data in the relevant domain, data-centric machine learning algorithms perform well. Some of the notable applications of AI in the health care industry are disease identification, outbreak prediction, drug development, smart health records, and so on. In this paper, we have investigated the role of ML in liver disease identification. The human liver is regarded as the most fundamentally functioning body organ, aiding in the digestion of food and removing poisonous compounds. A later diagnosis of liver illness will result in a life-threatening condition that will shorten a person's life. Machine learning can be used to fully assess the aspects of liver disease and forecast its severity. We used ML-based models KNN and SVM to identify liver disease in patients by selecting relevant features from the Indian liver patient dataset. Classification has been done by tuning different hyper parameters of the models, and some useful conclusions have been drawn at last.

**Keywords** Liver disease · Machine learning · KNN · SVM

## 1 Introduction

In the health care industry, a variety of basic and sophisticated machine learning and deep learning-based models are used. Located deep inside the body, the liver performs

---

G. Datta (✉) · A. Singh
School of Computer Science, University of Petroleum and Energy Studies, Dehradun, India
e-mail: gdatta1@yahoo.com

A. Singh
e-mail: anupam.singh@ddn.upes.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_11

many vital duties, the most significant of which are the removal of waste produced by our system, the digestion of food, and the maintenance of vitamin and energy stores. A problem with the liver may result in a range of serious conditions, including liver cancer. Early detection and treatment of these potentially fatal disorders are essential in order to reduce the risk of infection. Due to the high expense and complexity of diagnosing liver disease, a number of research have been done to classify individuals with liver disease using machine learning (ML) algorithms.

Primary liver cancer originates in the liver's cells. These malignant cells might form a single lump or begin in several locations throughout the liver. Patients with extensive liver impairment are more likely to develop several growth sites. Hepatocellular carcinoma (HCC), Cholangiocarcinoma (CR), Angiosarcoma, and Hepatoblastoma are the four types of primary liver cancer. Intelligent machine learning (ML) and neural networks (NN) models have played an important role in the diagnosis of liver disease. The goal of all of these algorithms has been to assess liver disease data and forecast sickness. In the medical profession, datasets have played an important role in illness analysis and detection. Because of the large and complicated data for analysts, illness diagnosis might be difficult at times. In the health care sector soft computing-based fuzzy logic techniques are used to control glucose level of diabetic patients [1]. Several previous research discovered that ML approaches provide a diverse set of tools, methodologies, and challenges to handle health care issues. ML algorithms are widely used in various applications in the health care sector. Proper hyper parameter tuning plays a crucial role to achieve greater accuracy. ML algorithms have wider range of hyper parameters and optimization of these parameters is very important [2]. ML algorithms are data-centric and requires huge data set to train such models. In health care sector one of the major challenge is lack of sufficient data set. For instance in lung cancer patient, sufficient data is not available. Hence Generative Adversarial Network (GAN) is widely used to generate synthetic data to solve this problem [3, 4]. Various deep learning algorithms such as convolutional neural networks are used for chest disease classification [5].

The computer-based technique is necessary for the non-invasive identification of asymptomatic, progressing, and potentially deadly chronic liver disorders.

There is another important use case of NLP, e.g. Machine Translation (MT) system which can be used in the health care sector, where there is a language barrier. Patients can speak in his native language about the symptoms of his disease and the doctor can understand in his language may be in English. M.T systems has passed through various phases, e.g. Rule-based, Statistical and these days we have Neural-based Translation Systems, i.e. Neural Machine Translation (NMT) system. NMT systems exploits the various machine learning and deep learning-based approach to enhance its performance. Performance of MT system in health care sector is very important to avoid any wrong communication between patients and doctors during diagnosis of a disease. Researchers are active to improve and evaluate the performance of MT systems [6–8].

Apart from disease predictions and other applications in the health care sector, ML model such as support vector machine (SVM)-based classifier helps in drug discovery process. SVM classifiers help classifying binary data, i.e. between drugs and non-drugs, based on its chemical composition and other properties. Even decision tree-based classifiers are also helpful in drug discovery process. Ensemble-based methods such random forests are also widely used in the health care industry such as in drug discovery process. Various deep learning-based models such as convolutional neural networks are widely used in medical imaging applications. Neural network-based models such as feed forward neural network, convolutional neural network, and recurrent neural networks have their own applications in various health care sector. Non-parametric-based classifiers such as KNN is widely used in various disease prediction. KNN failed to perform well when the dimensionality of data is too high. Different ML algorithms such as ANN, KNN, decision tree are widely used in breast cancer diagnosis.

Predictive medicine is a branch of medicine in which the threat of infection is predicted and precautionary measures are taken to either completely avoid the illness or significantly lessen its consequences on the individual. Some of the examples of predictive medicine are carrier testing, diagnostic testing, etc. Carrier testing is used to identify people who have one variant of a gene flaw that if present in both versions, causes a hereditary disorder. Diagnostic research is conducted to aid in the diagnosis and classification of a certain ailment. Even these days models based on machine learning are commonly utilized in genetics and genomics. Neural networks and deep learning models are widely used for accurate diagnosis purpose.

In this paper we have used non-parametric-based KNN and support vector machine (SVM) algorithms to predict liver disease. Tests for liver damage include aspartate aminotransferase, direct bilirubin, alkaline phosphate, alanine amino transfer, albumin, albumin, and the albumin-to-globulin ratio, to name a few. We have measured the accuracy of our models and have drawn some useful conclusion.

Rest of the paper is presented as follows. Section 2 describes some related research work in this field. Section 3 describes our methodology and experimentation with models in brief. Section 4 contains results and discussion. Finally, Sect. 5 includes some useful conclusion drawn.

## 2 Related Works

In the paper Yang et al. [9] authors obtained medical data from 1256 cirrhotic patients and used pre-processing to identify 81 characteristics from the irregular data. They examined several classification approaches to predict Hepatic Encephalopathy (HE) in cirrhotic patients using logistic regression, weighted random forest (WRF), SVM, and weighted SVM (WSVM). Authors also employed an additional 722 cirrhotic cases for external validation of the model.

Khan et al. [10] did a study which was primarily concerned with three modes of image acquisition, viz., ultrasonography, computed tomography, and magnetic resonance imaging. The authors gave an in-depth study with advantages and disadvantages for each pre-processing, attribute analysis, and classification approach used to fulfil clinical diagnostic tasks. However, a comparison of the research revealed that deep learning-based convolutional neural networks outperformed all other types of neural networks in terms of performance.

In the paper Nanyue et al. [11] by analysing the pulse signals of fatty liver patients, cirrhosis patients, and healthy persons there has been a significant difference noticed in the pulse signal between healthy person and fatty liver and cirrhosis patients. Researchers built a method by combining supervised and unsupervised ML algorithms.

In the paper, Terlapu et al. [12] authors used various ML algorithms such as logistic regression, Naïve Bayes, KNN, SVM, and decision tree. They got maximum accuracy (97.12%) with decision tree only as per their claim with liver disease dataset.

Vinutha et al. [13] used regression trees in their research to make predictions based on data from the liver disease dataset. Attributes including gender, weight, age, alcohol intake, platelet count, and albumin/globulin levels are included in the liver cirrhosis data set. The Mean Absolute Error (MAE), the Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are figured out. Liver illness is more likely to impact women than men, according to their model. Females are more likely to suffer from liver cirrhosis than men, according to the findings of the study.

Goldman et al. in their paper said [14] non-alcoholic fatty liver disease is most common disease amongst the population which spreads around 20–30% of the entire population. This disease is later on the cause of other cardiovascular disease and may lead to cirrhosis. Data from various sources fed to different machine learning algorithms to forecast individual risk. After careful data preparation and selecting various independent variables the models outperformed the existing techniques.

## 3 Methodology and Experimentation

We have followed the following steps to implement our work. We have collected data from kaggle Indian liver patient UCI [15].

Data pre-processing: Our data set has total records of 583. Out of which 416 liver patients and 167 non-liver patients' records. Firstly we thoroughly checked the data and clean the data to make it user ready. We have checked for if there is any missing value in the data or not. The important attributes of our data set is shown in Table 1. Based on these attributes the models classify if the person suffers from liver disease or not.

Rescale and splitted the data into train and test set. We then built our KNN model. $K$ is a hyper parameter in KNN model. In KNN we started with $K = 3$ and increased the $K$ value subsequently, i.e. $K = 5$ and so on. Model converges at $K = 54$ with training accuracy 74.55 and testing accuracy 71.26. In KNN algorithm $K$ value is the hyper parameter.

Making the best choice for the $k$ value is a critical step in developing a reasonable and accurate KNN model for a given situation. Models with a small $k$ value, fail to generalize successfully because their parameters are too narrowly defined. It's also quite sensitive to background noise. On the training set, the model has a high degree

**Table 1** Attributes/features of our data set

| Feature/attribute name | Description |
|---|---|
| Tot_Bil | Total bilirubin |
| Tot_Prot | Total proteins |
| Alamine | Alanine aminotransferase |
| Aspartate | Aspartate aminotransferase |
| A_G_Ratio | Albumin and globulin ratio |
| Albumin | Albumin |
| Dir_Bil | Direct bilirubin |
| Alk_Phos | Alkaline phosphatase |
| Age | Age |
| Sex | Gender |

of accuracy, but it will be a poor predictor of new data points. The result is likely to be an over fit model. If *k* is set to a value that is too big, the model becomes too generic and becomes incapable of reliably predicting data points in both the train and test sets. This is referred to as under fitting. The KNN classifier decides which class a data point should be based on the majority voting method. Suppose there are 54 closest points where *k* is set to 54. The classes of the 54 closest points are checked. Predictions are made based on the majority of the class. Similarly KNN regression takes the average of the 54 points that are the closest to each other, just like that.

Next we carried out our experimentation with SVM model. It is also a non-parametric classifier. Support vector machines (SVMs) are a kind of supervised learning technique that may be used to solve classification and regression issues in a variety of situations. In order to make data points linearly separable, SVM algorithms translate them from a low-dimensional to a high-dimensional space. A hyperplane is then constructed as the classification border to split data points into categories. As a result, the kernel type would be an extremely important hyper parameter to modify. An essential hyper parameter in SVM is the kernel function, which is used to quantify the similarity between two data points and is one of the most critical hyper parameters that must be tuned. The linear kernel, the radial basis function (RBF), the polynomial kernel, and the sigmoid kernel are all common kinds of kernels in SVM. We have used RBF kernel in our model and got the accuracy 66.43. Figure 1 depicts a visual illustration of the procedures involved in the construction of a standard ML model.

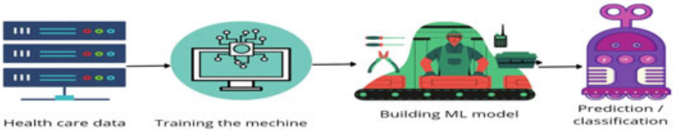


**Fig. 1** ML model building typical steps

## 4 Results and Discussion

Analysing the results of both the models we can see that in case of KNN model, when the value of $K$ is too small then the model tends to over fit and whilst the $K$ value is too large the model tends to behave like under fitting. As described before we started with $K = 3$ then $K = 5$ and got the accuracy 63.01 and 63.91, respectively. Ideally $K$ value needs to be selected on the basis of square root of total training data available. Hence, proper hyper parameter selection is extremely important. In support vector machine we have selected RBF as one of kernel which is the hyper parameter in SVM. The accuracy of both the models with their respective hyper parameters are reported in Table 2.

The graphical representation of KNN and its behaviour with respect to different $K$ values is shown in Fig. 2.

In a separate graph we have presented different values of $K$ viz. 3, 5, and 54 and the improvement in accuracy values (Fig. 3). After certain $K$ towards upper side value model becomes under fit with declining values of both training and testing accuracies.

The comparative analysis of these two models is shown in Fig. 4. For KNN with selected set of hyper parameters the maximum accuracy achieved is 74.55 whereas with SVM with its kernel (= rbf) is 66.43.

**Table 2** Models' outputs with their corresponding hyper parameters

| Accuracy | Classifier | List of hyper parameters |
|---|---|---|
| 74.55 | KNN | Metric parameter = Minkowsky, nearest neighbour = 54, $p = 2$ |
| 63.91 | KNN | Metric parameter = Minkowsky, nearest neighbour = 5, $p = 2$ |
| 63.01 | KNN | Metric parameter = Minkowsky, nearest neighbour = 3, $p = 2$ |
| 66.43 | SVM | Kernel = rbf |

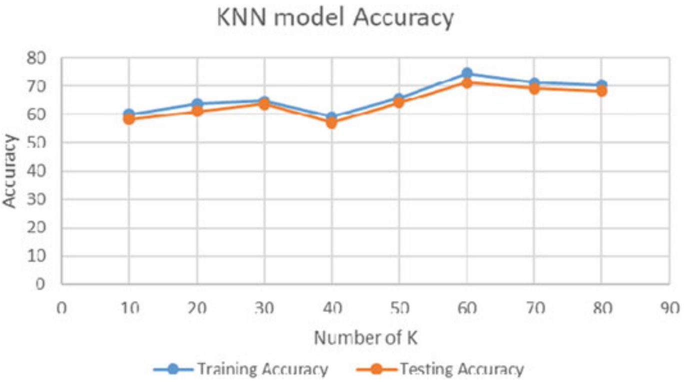


**Fig. 2** Graphical representation of training and testing accuracy of KNN with different $K$ values

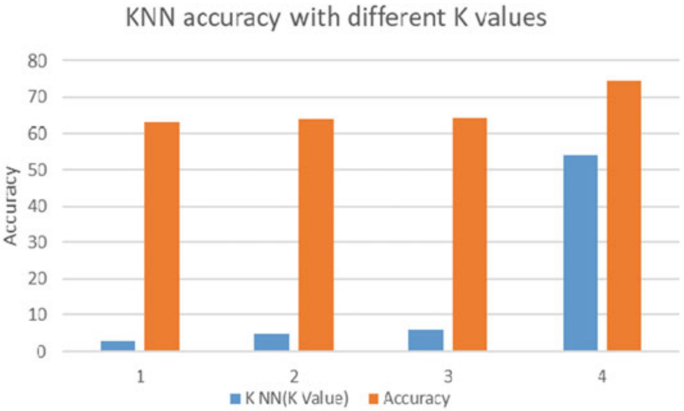


**Fig. 3** KNN accuracy with $K$ values ranging from $K = 3$, $K = 5$, and $K = 54$

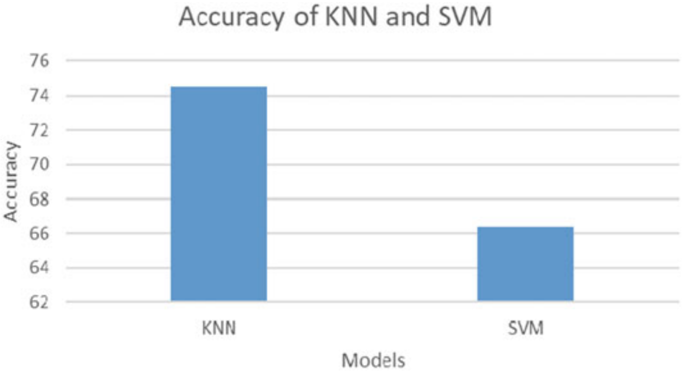


**Fig. 4** Comparison of two models in terms of their accuracy and selected hyper parameters

## 5 Conclusion

Although ML algorithms are widely used in solving various problems in health care and other sectors these days, proper hyper parameter optimization/selection, correct model selection as per the data set, is very important to achieve greater accuracy. Interpretability of ML algorithms is a great challenge that we face these days. Hence, we can't blindly believe the results of ML algorithms. ML algorithms are data-centric and sometimes the possibility of bias in data can't be ruled out. One needs to carefully pre-process and understand the data before training the model to achieve a reliable and good result.

## References

1. Zhu KY, Liu WD, Xiao Y (2014) Application of fuzzy logic control for regulation of glucose level of diabetic patient. Intell Syst Ref Libr 56:47–64. https://doi.org/10.1007/978-3-642-40017-9_3
2. Yang L, Shami A (2020) On hyperparameter optimization of machine learning algorithms: theory and practice. Neurocomputing 415:295–316. https://doi.org/10.1016/j.neucom.2020.07.061
3. Lin CH, Lin CJ, Li YC, Wang SH (2021) Using generative adversarial networks and parameter optimization of convolutional neural networks for lung tumor classification. Appl Sci (Switz) 11(2):1–17. https://doi.org/10.3390/app11020480
4. Gonzalez-Abril L, Angulo C, Ortega JA, Lopez-Guerra JL (2021) Generative adversarial networks for anonymized healthcare of lung cancer patients. Electronics 10(18). https://doi.org/10.3390/ELECTRONICS10182220
5. Ranjan R, Singh A, Rizvi A, Srivastava T (2020) Classification of chest diseases using convolutional neural network. In: Proceedings of first international conference on computing, communications, and cyber-security (IC4S 2019). Lecture notes in networks and systems, vol 121. Springer, Singapore. https://doi.org/10.1007/978-981-15-3369-3_18
6. Datta G, Joshi N, Gupta K (2021) Empirical analysis of performance of MT systems and its metrics for english to Bengali: a black box-based approach. In: Intelligent systems, technologies and applications. Advances in intelligent systems and computing, vol 1353. Springer, Singapore. https://doi.org/10.1007/978-981-16-0730-1_24
7. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K (2018) Development of machine translation technology for assisting health communication: a systematic review. J Biomed Inform 85:56–67. https://doi.org/10.1016/j.jbi.2018.07.018
8. Costa-Jussà MR, Farrús M, Pons JS (2012) Machine translation in medicine: a quality analysis of statistical machine translation in the medical domain. In: Conference on advanced research in scientific areas, pp 1995–1998. Retrieved from http://www.arsa-conf.com/archive/?q=2012-ARSA-156::2
9. Yang H et al (2021) Using machine learning methods to predict hepatic encephalopathy in cirrhotic patients with unbalanced data. Comput Methods Programs Biomed 211:106420. https://doi.org/10.1016/j.cmpb.2021.106420
10. Khan RA, Luo Y, Wu F-X (2022) Machine learning based liver disease diagnosis: a systematic review. Neurocomputing 468:492–509. https://doi.org/10.1016/j.neucom.2021.08.138
11. Nanyue W et al. (2015) Pulse diagnosis signals analysis of fatty liver disease and cirrhosis patients by using machine learning. Sci World J 2015. https://doi.org/10.1155/2015/859192
12. Terlapu PV, Prasad R, Sadi R, Pondreti RK (2021) Intelligent identification of liver diseases (IILD) based on incremental hidden layer neurons ANN model. https://doi.org/10.1007/978-981-15-8439-8
13. Vinutha MR, Chandrika J (2021) Prediction of liver disease using regression tree. Int J Online Biomed Eng 17(2):164–172. https://doi.org/10.3991/ijoe.v17i02.19287
14. Goldman O et al. (2021) Non-alcoholic Fatty liver and liver fibrosis predictive analytics: risk prediction and machine learning techniques for improved preventive medicine. J Med Syst 45(2). https://doi.org/10.1007/s10916-020-01693-5
15. https://www.kaggle.com/uciml/indian-liver-patient-records?select=indian_liver_patient.csv

# Telemedicine: Enhancing and Reimagining the Public Health in India

**Shweta Pandey, Abhishek Jha, Anjali Jha, Anurag Lokhande, Zahara Khan, Ulhas W. Karkhanis, and Aradhana Goutam**

**Abstract** Concerning the movement of patients in regards to the COVID-19 pandemic, consulting the physician face to face is difficult. Due to this, it warranted in adopting the culture of telemedicine in urgency. The COVID-19 has majority adopted in expanding the telemedicine culture in the healthcare sector. This study's goals were to assess patient satisfaction, bridge the physician–patient communication gap, and assess the effectiveness and safety of telemedicine versus in-person consultations for patients. There is a rise in the telemedicine sector as different segments and socialization is involved which in the future will expand to new business models. The study was carried on 100 patients of sample size of broad age group. The motive of this paper is to determine the challenges and preferences in India to implement

---

S. Pandey (✉) · A. Jha · A. Jha · A. Lokhande · Z. Khan
PGDM-PM, IES Management College and Research Centre, Mumbai Vishwakarma
M D Lotlikar Vidya Sankul, 791, S K Marg Bandra Reclaimation, Mumbai 400050, India
e-mail: shweta.pandey21@ies.edu

A. Jha
e-mail: abhishek.jha21@ies.edu

A. Jha
e-mail: anjali.jha21@ies.edu

A. Lokhande
e-mail: anurag.lokhande21@ies.edu

Z. Khan
e-mail: zahara.khan21@ies.edu

U. W. Karkhanis
Pharma
Management, IES Management College and Research Centre, Mumbai Vishwakarma M D
Lotlikar Vidya Sankul, 791, S K Marg, Bandra Reclaimation, Mumbai 400050, India
e-mail: ulhas.karkhanis@ies.edu

A. Goutam
Department of IT, IES Management College and Research Centre, Mumbai Vishwakarma M D
Lotlikar Vidya Sankul, 791, S K Marg, Bandra Reclaimation, Mumbai 400050, India
e-mail: aradhana.goutam@ies.edu

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_12

this emerging technology of telemedicine across. In conclusion, there is advancement in medical field administrative where post-pandemic still many people prefer tele consultation.

**Keywords** Telemedicine · Patient satisfaction · Health outcome

## 1 Introduction

"Telemedicine is one of the productive fields in India where the private sector has taken the lead and participated actively in the administration of public health. Among the telemedicine companies operating in the Indian private sector are Narayana Hrudayalaya, Apollo Telemedicine Enterprises, Asia Heart Foundation, Escorts Heart Institute, Amrita Institute of Medical Sciences, and Aravind Eye Care. Practically 75% of the population lives in country regions lacking admittance to clinical aptitude and foundation. Besides, medical services conveyance is troublesome in the unfriendly geological landscape, for example, mountain locales in the upper east, deserts of the northwest, and the seaward islands of Andaman and Lakshadweep. The capability of telemedicine technology in giving medical services admittance to country populaces and far-flung regions has for quite some time been understood, and numerous specialized services of the Government of India like Information Technology, Science and Innovation, and Space have been exploring different avenues regarding telemedicine pilot projects since mid-2000. In view of the effective result of these pilots, the Ministry of Health and Family Welfare has now taken on telemedicine into the National Rural Health Mission, and drive zeroed in on progress of the provincial medical services conveyance framework. All over the country, a few telemedicine drives have been taken up by both government and private area associations with government and state financing. Some have taken on a couple of modules into their well-being framework" [1].

## 2 Telemedicine in India

Both, popular society and personal regions were related to starting some telemedicine programs in India. Curiously, the underlying telemedicine framework in India turned into deliberate and spearheaded through a private mission, Apollo clinics, as a crew with Indian Space Research Organization (ISRO) at a metropolis known as Aragonda in Andhra Pradesh. Therefore, with the assist of Department of information Technology (DIT), ISRO, Department of health and Family Welfare, Ministry of External Affairs (MEA), National Informatics Center (NIC), and separate country legislatures, the business enterprise of telemedicine prolonged the country over. The chief showing businesses move about as the focal center factors, to be unique, AIIMA, New Delhi; PGIMER, Chandigarh; SGPGIMS, Lucknow; TMH, Mumbai; PBDSPGIMS, Rohtak; and CSMMU, Lucknow. A few company scientific organizations have additionally installation telemedicine focuses to expand their administrations: Amrita Institute of Medical Sciences (Kochi), Narayana Hrudayalaya (Bangalore), AECS (Madurai), Apollo hospitals (Hyderabad), Sankara Nethralaya (Chennai), and Sri Ramachandra Medical Center (Chennai). Steadily, the telemedicine community has stretched out the country over with extra than 400 telemedicine ranges in India. ISRO's telemedicine program interfaces 245 emergency clinics, 205 locale/rustic clinics, and 40 outstanding-area of expertise clinics. In any case, the sizeable majority of those drives stay bound to a "pilot project" degree and reception of telemedicine in to well-known hospital therapy conveyance frameworks remains intricate.

Province public telemedicine network The Federal Ministry of Health and Family Welfare has also acknowledged that telemedicine has a favorable impact on the country's continued advancement of medical care. It has chosen to set up a public country telemedicine network. In this undertaking, a public organization for interlinking every single clinical school/University Hospitals the nation over have been proposed. The National Medical College Network is an organization that promotes clinical education and e-health initiatives across the nation. Under the National Knowledge Network Project, 150 clinical schools across the country are currently connected by a quick fiber-optic-based Internet. The National Rural Telemedicine Network is working to connect rural outlying medical care populations with local clinics, tertiary care facilities, and academic medical clinics in the field of medical care delivery [2].

"Association of private venture in the field of telemedicine in IndThe privateate area/medical services industry has additionally perceived the capability of telemedicine. Through the effective execution of a few pilot projects including the public authority and private undertakings, telemedicine has been laid out as a strong and solid medical services conveyance framework in India. The Apollo Telemedicine Networking Foundation, a not-for-benefit association and a piece of the Apollo Hospitals Group has been endeavor telemedicine projects. Under the administration of Dr. Ganapathy, a noticeable neurosurgeon and generally viewed as the dad of Indian Telemedicine, Apollo Hospitals worked together with ISRO to set up India's first Rural Telemedicine Center at Aragonda, a little town in the State of Andhra Pradesh. The world's first VSAT empowered, present day optional consideration medical clinic was laid out on March 24th 2000. From that point forward Aragonda Apollo Telemedicine focus and Hospital has been the foundation contextual analysis model for the whole telemedicine medical services industry. Today ATNF has arisen as India's single biggest private medical care supplier in the space of Telemedicine with north of 125 fringe places in India and 10 abroad" [2].

## 3 Telemedicine Guidelines During COVID-19

The Board of Governors of (MCI), i.e., the Medical Council of India, the most recent clinical training supervisor in India, who pre-arranged the rules in a conversation with the principal organizing body the National Institution for Transforming India (NITI Aayog), has made an effort to fill a significant gap: absence of regulation and a structure for the moral act of telemedicine. The standards list video, sound, and text as three strategies for resemblance and map out the game plans for their usage by specialist, including limits. The principles need clarity about assurance and data to use, for patients and specialists. They set the onus by and large on experts to stay aware of records of all exchanges of correspondence among themselves and patients. The principles don't yet decide the term for taking care of data nor limiting to extra use of this data. The instructions essentially require the master to be familiar with the data security and assurance guidelines and adhere to them. Security concerns arise as nuances, including a patient's area, and other reasonable unmistakable confirmation, are relied upon to be recorded by trained professional. The standards explain the thoughts of inferred and express consent anyway basic beginning of a telemedicine interview by an individual is considered as suggested consent. The guidelines need to explain more on consent in a tele-consultation and approaches to getting and recording it. The standards moreover come up short on instruments for settling the protests of patients or specialists.

**Objectives**

1. To determine the patient's satisfaction difference between video and person visits.
2. To understand the opportunities and challenges in adapting via telehealth services.
3. To bridge the gap between patients and physicians.

## 4 Literature Review

According to the diary of general well-being strategy, foundations funded by the Union government and those backed by the State immediately provided services by electronic means. Soon after the end of emergency clinic units, the Ministry of health and Family Welfare laid out the country's telemedicine strategy rules. The Medical Council of India (MCI) Board of Governors, the most recent clinical training supervisor in India who in advance with the primary organizing body, the NITI Aayog (National Institution for Transforming India), has made an effort to close a sizable gap: the lack of guidelines and a framework for the moral act of telemedicine. The regulations outline the conditions for their use by experts, including any restrictions, and mention video, sound, and text as three possible forms of correspondence [3].

"A portion of the regulations connecting with the clinical calling incorporate the Medications and Cosmetics Act, 1940, and Drugs and Cosmetics Rules, 1945, the Indian Medical Council Act, 1956, the Indian Medical Chamber (Professional lead, Etiquette and Ethics) Regulations, 2002, and the Clinical Establishments (Registration and Guideline) Act, 2010 ('Clinical Establishments Act'). The regulations connected with ICT incorporate the Information Technology Act, 2000 (IT Act), 18 the Rules, 2011, 19 of Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Data) the Information Technology (Mediators Guidelines) Rules, 2011, 20 Unsolicited Commercial Correspondences Regulations, 2007, 21 and Telecom Commercial Correspondence Customer Preference Regulations, 2010 ('TCCP Regulations'). The Information Technology (IT) Act contains arrangements protecting the security and protection of data traded utilizing implies perceived under the IT Act. All organizations enlisted under innovation or giving innovation administrations are represented by the IT Act" [4].

Government associations like the Indian Space Research Organization (ISRO), the Department of Informational Technology (DIT), the MOHFW, and the Ministry of External Affairs all played critical roles in the development of telemedicine services in India. These offices have set up telemedicine connections between several rural areas and important emergency clinics in urban areas. Certain patient's in India can find the use of telemedicine a bit difficult, particularly old and uninformed patients who are coming up short on specialized abilities for exploring the above discussion stages [5].

The rules outlined by MOHFW have been distributed under the Indian Medical Council (IMC) Act. RMPs who might need to rehearse telemedicine ought to submit to these rules and need to finish a required Internet-based course within 3 years of its warning. The rules given by MOHFW, propose a video method of correspondence for the first counsel anyway as per Diabetes and Metabolic Syndrome: Clinical Research and reviews recommends that the first counsel for patients with diabetes ought to preferably be up close and personal on the off chance that conceivable in light of the fact that actual assessment can't be supplanted by telemedicine except if one more medical services specialist in actual vicinity to the patient can transfer assessment discoveries by means of telemedicine stages. During the hour of Coronavirus, the MOHFW rules, just a state-enrolled clinical expert (RMP) is permitted to rehearse telemedicine in India. Further, the rules given by MOHFW recommend a video method of correspondence for the first counsel in any case, as indicated by Diabetes and Metabolic syndrome: Clinical Research and Reviews proposed that the first counsel ought to be up close and personal on the grounds that actual assessment can't be supplanted by telemedicine [5].

Far off locales all around the planet face a few obstructions in giving available, sensible, and pinnacle-notch medical care administrations. These barriers incorporate financial necessities, complex geographic areas, unlucky foundations, and a lack of prepared medical care professionals in the district. Telemedicine, or the conveyance of hospital therapy administrations through statistics and correspondence advances (ICT), can possibly reduce variations in hospital therapy, further broaden wellness results, and supply training to scientific services suppliers. Ongoing advances in innovation, as an instance, the utilization of cell telephones, in addition developed Internet framework, digitalized statistics, and digital scientific facts, have made telemedicine attainable. Telemedicine may be applied in an assortment of methods, and the kind of telemedicine apparatuses applied is based vigorously upon the necessities and assets of a populace. Creators coordinated this audit by using gross public pay (GNI) in step with capita in US bucks, a World Bank order, to introduce exceptional telemedicine drives almost about the United States' financial status.

Akiyama and Yoo7 revealed that teledermatology and teleradiology frameworks confirmed promising results in monetary proficiency. As one of the maximum evolved economies in Asia with an excessive limit with recognize to ICT, Singapore has had some teledermatology drives tested. A Web-based teledermatology framework changed into assessed amongst sufferers with nonurgent pores and skin situations in a nursing domestic. Seghers and colleagues 11 directed a bit impending review on the Institute of Mental Health in Singapore to comparison teledermatology by means of regular videoconferencing and an up-near and personal convention in a meeting of standardized intellectual sufferers with continual skin illnesses. They closed 24 Jha et al. that teledermatology turned into quite a lot as successful as eye interview in regards to demonstrative precision and the executive's consequences in light of a simple diploma of association between the forms of meetings.

As the second maximum populated USA on the planet, India has been using e-health to arrive at its extra weak residents for a totally long time. In 2008, the World Health Partners' (WHP) Sky Program installation telemedicine middle points in provincial areas in India. "SkyHealth" facilities are institutions run with the aid of neighborhood commercial enterprise visionaries that deliver varying media correspondence between sufferers in provincial towns and metropolitan doctors. Over 65% of the population in Pakistan lives in provincial regions and countenances geographic difficulties whilst attending to reasonable and exceptional wellness care. 34 to warfare the lack of hospital therapy administrations there, Telenor, a large media communications corporation adjusting nations across Asia, sent off a telemedicine management in 2008 after their accomplishment in Bangladesh. C.H.E.S.T has successfully arrived at networks in remote and underestimated areas of Nepal and proceeds to effectively deal with this system.

The report gives the build yearly development rate (CAGR) for the worldwide Telemedicine market during the conjecture time frame. The report includes a detailed executive summary as well as a sneak peek at how several pieces that were remembered for the duration of the assessment are developing. The research also provides insight into the key features of the global Telemedicine market that are changing. These records fill in as important devices for existing business sector players as well concerning elements keen on entering the worldwide Telemedicine market [6].

The report dives into the serious scene of the worldwide Telemedicine market. Central participants working in the worldwide Telemedicine market have been recognized. Organization outline, monetary standings, late turns of events, and SWOT are a portion of the traits of players in the worldwide Telemedicine market that have been profiled in the report [6].

The move united states lockdown and resultant limitations further augmented the all-round current essential unnoticed remedy need for substance use problems (SUD). Telemedicine may want to expect a vast component in lessening the remedy hollow via working on the entrance and accessibility of proof primarily based attention, during this season of emergency as well as in regular instances. A review has shown the viability of tele-consultation in a joint attempt with the crucial attention health practitioner for offering mental types of assistance to vagrants in a recuperation vicinity.

The Telemedicine Practice Guideline (TPG) changed into introduced on March 25th 2020 and became formally informed in the Gazette of the Government of India on May 14th 2020 Telepsychiatry rules were additionally brought in May 2020. The point of each the guidelines changed into to offer pragmatic steerage to clinicians with respect to the 'setting up, execution, corporation, and arrangement of tele-consultation management so it tends to be included into the everyday medical practice.

A more extensive use of cell phone-based tele-consultation for patients with substance use problems began during the lockdown because of the COVID-19 pandemic. In the review the telemedicine model embraced by the Drug Deaddiction and Treatment Center of the Postgraduate Institute of Medical Education and Research, Chandigarh was shown which have a multi-disciplinary out-patient help comprising of specialists, clinical specialists, attendants, clinical therapists, and guides. The test was to consolidate and coordinate the act of telemedicine inside the TPG structure yet in addition in a stepwise progressive operations administration model contingent on the need of the circumstance [7].

As indicated with the aid of our assessment, the administrations gave by means of telemedicine should have comparable requirements of professional duty and protection as preferred conferences. Specifically, tele-consultation should be robotically paid, with recommend and answers joined into a traditional report sent with the aid of e-mail to the affected person, along with a receipt. The specialists undermined the arrival of guidance and answers exclusively through verbal correspondence or straightforwardly through communication, albeit the closing preference would anyways have its lawful worth.

Tele-consultation was rather considered proper for the conclusion and the executives of pilonidal illness (middle worth, 7), probably in light of the fact that the analysis is worked with by the utilization of top quality pictures, as well as by the run of the mill limitation of the sickness. The significant job of actual assessment may again clarify this result as featured in a new study of patient and supplier experiences. 25 interestingly, given the assessed predominance of hemorrhoidal sickness (i.e., 4.4%) 26 and the rate of butt-centric crevice (i.e., 0.11 individual years, bringing about normal lifetime hazard of 7.8%), 27 the conjunction of the two circumstances isn't insignificant and must be avoided on an ordinary visit [3].

Taking the whole thing into account, what must the reasoning of telehealth be? Maybe equal to that of two decades old. "Information improvement need to be given a principal element with inside the replace of the scientific advantages shape awaiting an important nice development is to be achieved". Recognizing its endpoints and defensive patients' prosperity is prime for this work [3].

Patients were isolated into a telemedicine accomplice, with preoperative conference and careful navigation directed through telemedicine, and a conventional in-person companion. 28 had their preoperative consideration solely directed through telemedicine and were in the telemedicine accomplice. Sixty patients were preoperatively found face to face, 28 through telemedicine, and six had their preoperative visit changed over from telemedicine to face to face an assortment of reason. Concerning follow-up, most patient selected telemedicine visits, with 61 patients seen through telemedicine, 30 patients found face to face, and three had their postoperative visit changed over from telemedicine to face to face. Patients were enlisted from March 2020 until October 2020, the period in which our establishment's compulsory COVID-19 telemedicine convention was active by which all new patients were offered an underlying telemedicine conference yet were given choices for in-persons arrangement in response to popular demand. The telemedicine partner contained patients who had preoperative assessment only led by means of telemedicine and continued straightforwardly to a medical procedure [8].

According to McCool RR, Davies L. Where does telemedicine fit into otolaryngology? An appraisal of telemedicine qualification among otolaryngology diagnoses. Recent surveys of the expense adequacy of telemedicine drives found that telemedicine can lessen the expense of medical care conveyance, accordingly making medical care administrations reasonable and available. To make telemedicine a fundamental piece of the biological system of medical care conveyance, and to assure dynamic cooperation of the private area on a practical fact, a very less arrangement drives are expected for assuring the monetary suitability of telemedicine projects, for example, lucidity on recompense of therapy gave by telemedicine. Defeating the common shortage of paramedic faculty and preparing the staff in the utilization of telemedicine is vital to expanding the reception of telemedicine all over provincial regions [8].

Defeating the socio-social hindrances to stretch out telemedicine reception to country regions, seemingly, relies on the fuse of telemedicine in essential consideration. The granular perspective to execute telemedicine at the essential consideration level would, ostensibly, improve the viability and acknowledgment of telemedicine at the auxiliary and tertiary levels [8].

"This could be the future for medical services in India, where both the general population and the private well-being areas are giving significant measures of cash and work to standard telemedicine. Defenders of telemedicine consider it to be an answer for a portion of the lacks in the country's well-being area. India faces a lack of specialists, especially in rustic regions, where almost 70% of the populace lives. Specialists and clinics are generally packed in urban communities, and as a result, medical services in provincial India is lacking or missing. One of the principle points of telemedicine is to give medical services to India's most unfortunate individuals. Since then-US President Bill Clinton launched India's first telemedicine initiative at Apollo Aragonda Hospital in the state of Andhra Pradesh in 2000, a lot has changed. Today, there are around 500 telemedicine focuses connected with around 50 expert emergency clinics the nation over. In the next few years this figure is expected to increase. The public authority's eleventh Five-Year-Plan (2007–2012) assigned 2000 million rupees (about US$50 million) to telemedicine. (2007–2012) allocated 2000 million rupees (about US$50 million) to telemedicine" [9].

## 5 Research Methodology

This research methodology has been utilized to conduct comprehensive research on the telemedicine market in India and to arrive at conclusions on its growth prospects amidst COVID-19, opportunities, and challenges for telehealth services. We had done primary and secondary research which has supported to draw of an accurate and reliable conclusion.

For primary research, a structured questionnaire was designed and administered to a group of 100 people. The survey was administered online, and only one response per person was taken into account. The research consisted of both qualitative and quantitative studies.

There were certain closed-ended questions to know people's approach toward telemedicine, are the people enough adaptable to this change in the healthcare system, then there were some questions related to the awareness of various existing telemedicine platform, the scope of improvement from the patient's perspective.

We had interviewed a few doctors as well which serves as a validation from the perspective of the doctors about this growing sector in healthcare.

## 6 Physicians Interview

The majority of the doctors that we had interviewed were MBBS graduates and very few did their specialization.

Questions that were asked during the interview:

(1) Do you prefer telemedicine?

Almost all of the physicians that we had interviewed preferred telemedicine and they told that it is the best facility for the patients who cannot reach out to the doctor personally and it has proven to be very useful in the situations like a pandemic.

(2) Pros and Cons of telemedicine?

**Pros**: Most of the physicians mentioned it as easily access time-saving for both doctors and the patient as a doctor could reach many patients in lesser time and for a patient it is money-saving as far off patients can get in touch with the doctor for small ailments without having to pay for the travel and a safer and convenient method for everyone in a pandemic situation.

**Cons**: In response to the cons of telemedicine, physicians mentioned different points which might be the possible reasons according to them leading to certain doubts in patient's mind while opting for telemedicine. These are inaccuracies in diagnosis because e detailed physical examination is not possible or because of lack of proper exposure, a handover technology is required which is still scarce in certain regions of India, also there are limited tools of assessment, Inadequate communication services.

(3) How did you with up in the pandemic with telemedicine?

For a few doctors, online instructions regarding the diagnostic and treatment strategies helped. Online CT reports also helped save time and lives, instructions regarding preventive care were a major plus point. Many doctors consulted many non-emergency patients through telemedicine which provided safety to patients as well as doctors. All the doctors were being cautious with symptoms and taking regular follow-ups with the patients. Doctors said patients with mild symptoms could be directed well and the physical contact was lessened. Answering patients' queries made them decide whether they need to get admitted to the hospital or stay at home.

(4) What is the impact of pandemic on your practice?

Almost all the doctors said more working hours resulted in increased stress. Increased chances of infection because of high contact with the positive patients, hectic schedule during the pandemic. Patients requiring surgical management reduced in hospitals only emergency patients were taken in the operation theater, many doctors reported that patients were preferring to discuss their symptoms over a call rather than a visit. Doctors found that there was a reduction in the overall variety of cases and an increase in pandemic-related patients.

For secondary research, we had gone through 10 research papers of a few Indian and International Journals where it includes research design, procedures, analysis, and sample collection which we intend to use in our research. In this research we have focused on open ended interview questions to identify the key terms in this study as we proposed.

## 7 Results Analysis

The research has been done with quantitative analysis methods to see answer to the research questionnaire.

An analysis feedback form consists of 10 no of questions developed, validated, and circulated among the 110 no of consumers (through random sampling with the help of direct interview questions, Google form, Email, WhatsApp, etc. To know the preferences toward tele-consultation.

In this study we have observed that the majority of the people, i.e., 74.5% are youth who preferred tele-consultation. About 40.6% of people preferred telemedicine over the face to face consultation. The majority of the people 60.4% are satisfied with the telemedicine consultation, whereas 39.6% of people were not satisfied with the telemedicine consultation.

71.7% of the people will recommend telemedicine to their friends and family members. 74.5% of the people who promote their college about telemedicine consultation.

Out of a total, 70.8% of people think telemedicine consultation is beneficial for their daily practice. 47.2% of the consumer have taken tele-consultation and when we

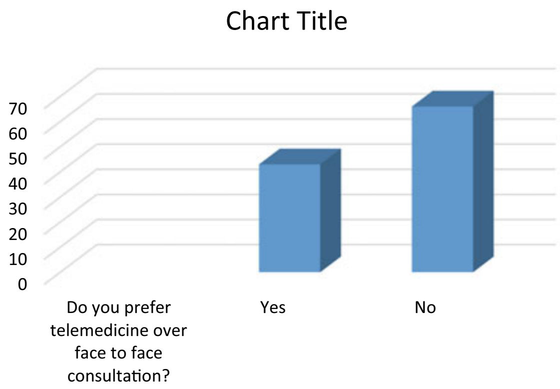


**Fig. 1** Preference of telemedicine

ask about the experience 38.4% of consumers have given a positive response toward tele-consultation. They believe that tele-consultation is seen as the futuristic platform for telemedicine. Later asked them whether the telemedicine sector will grow post-pandemic to which 83% shows positive response majority city of the people believe that tele-consultation is time saving, while 45.3% think that telemedicine is comfortable and convincing. 14.2% believe that it provides better patient care additionally 39.6% show fast assessment during telemedicine consultation.

When asked whether doctors are able to understand their health issues has have provided them with adequate attention. 76.4% of people agree that their doctor provided them with the utmost care and was able to understand their problems.

## 8 Result

See Figs. 1, 2, 3 and 4.

1. **Do you prefer telemedicine over face to face consultation?**
2. **Have you taken any tele-consultation?**
3. **How was your experience?**
4. **Your view on telemedicine?**

## 9 Conclusion

The conversation introduced in this paper proposes that few telemedicine drives have been sent off in India. It shows in general; most patients were happy with their tele-visit. Also, phone and video tele-visits were comparable with respect to patient

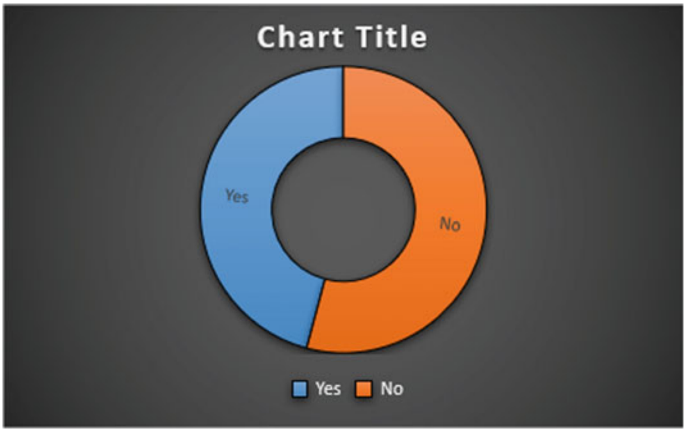


**Fig. 2** Tele-consultation preference

**Fig. 3** Experience on telemedicine

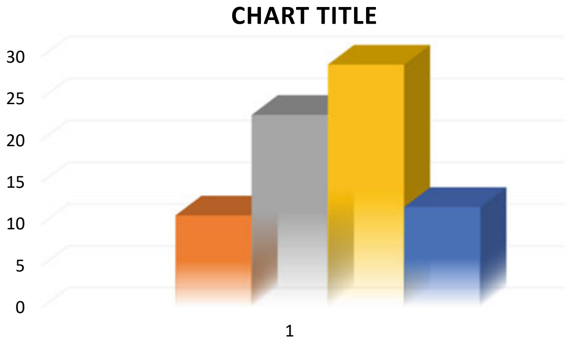
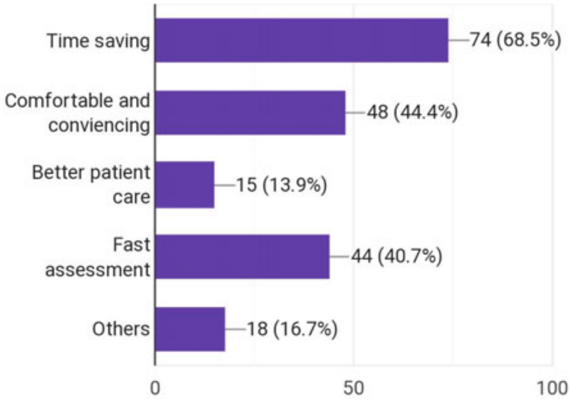


**Fig. 4** Views on telemedicine



feelings, patient attributes, and visit result. Endeavors to build access and inclusion of telehealth, especially phone tele-visits, should go on past the COVID-19 pandemic. Telehealth plays immediate and aberrant parts in lessening the spread of contaminations by empowering physical removing, following side effects and flare-ups, and supporting policymakers in expecting needs and choosing fitting and ideal intercessions. Imagining a telehealth-empowered future, one should recall that numerous clinical circumstances require actual assessments or intercessions and that care is a human and social action. Until this point in time, venture has been lacking in creating advancements that work for clinicians and patients and adjusting medical care frameworks and ways of life to completely take advantage of these innovations. The COVID-19 pandemic addresses a genuine open door and motivation to create progressed telehealth arrangements that can change medical care and individuals' lives, both locally and globally. Telemedicine can assist us with defeating the boundaries to getting to medical care benefits the–distance among supplier and patient, admittance to transportation, divided care caused because of holes in arrangements, and absence of medical care access. To finish up, while telemedicine might conceivably improve the openness, reasonableness, and nature of medical care administrations in India, as shown by a few pilot projects, effective and practical increasing of such drives requires tending to strategy/administrative, infrastructural, human asset, and socio-social issues.

## 10 Limitations

The limitation of this study in terms of data collection is due to the rationale of the COVID-19 pandemic, during which close physical contact is strictly prohibited and a social distancing policy is required to avoid a rapid increase in cases. As a result, gathering knowledge and knowledge will only be accomplished through the use of a web survey. Another limitation is that thanks to time constrain we could conduct only a little survey with a sample size of 100 and for the physician's interview we circulated our questionnaire to 10 doctors.

## References

1. Mishra SK, Kapoor L, Singh IP (2009) Telemedicine in India: current scenario and the future. Telemedicine e-Health 15(6):568–575. https://doi.org/10.1089/tmj.2009.0059
2. Mathur P et al. (2017) Evolving role of telemedicine in health care delivery in India. Prim Health Care Open Access 07(01). https://doi.org/10.4172/2167-1079.1000260
3. Gallo G et al (2021) E-consensus on telemedicine in proctology: a RAND/UCLA-modified study. Surgery 170(2):405–411. https://doi.org/10.1016/j.surg.2021.01.049
4. Ateriya N et al (2018) Telemedicine and virtual consultation: The Indian perspective. Natl Med J India 31(4):215. https://doi.org/10.4103/0970-258X.258220

5. Ghosh A, Gupta R, Misra A (2020) Telemedicine for diabetes care in India during COVID19 pandemic and national lockdown period: guidelines for physicians. Diabetes Metab Syndr 14(4):273–276. https://doi.org/10.1016/j.dsx.2020.04.001
6. ProQuestDocuments-2022-02-15.pdf
7. Ghosh A et al (2021) Telemedicine-assisted stepwise approach of service delivery for substance use disorders in India. Asian J Psychiatr 58:102582. https://doi.org/10.1016/j.ajp.2021.102582
8. Boles RW, Zheng M, Kwon D (2022) Expanded use of telemedicine for thyroid and parathyroid surgery in the COVID-19 era and beyond. Am J Otolaryngol 43(2):103393. https://doi.org/10.1016/j.amjoto.2022.103393
9. Solberg KE (2008) Telemedicine set to grow in India over the next 5 years. Lancet 371(9606):17–18. https://doi.org/10.1016/S0140-6736(08)60052-5

# Effect of Point of Service on Health Department Student's Creativity in Comprehensive Universities of Ethiopia: Moderating Role of Public-Private Partnership and Mediating Role of Work Place Learning

**Chala Wata Dereso, Shashi Kant, Maheswaran Muthuraman, and Gemechu Tufa**

**Abstract** As a referral plan, point of service customers pay less if use healthcare network. This concept is new for Ethiopia so education institutes have the responsibility to first educate their healthcare students with help of work place learning in public-private partnership environment. So that point of services help in enhancing student's efficiency by nurturing creativity. Therefore, present research used point of service as the antecedent and students creativity as the predicate. Public-private partnership was employed as a moderator to investigation work place learning associations with antecedent and result of consequence. Cross-sectional research design was used. Students, academic and non-academic staff of comprehensive universities in Ethiopia especially related with health services considered as the sample group. Curvilinear relationship between point of service and work place learning was measured with help of Confirmatory Factor Analysis (CFA). PROCESS macro for mediation, moderation and conditional process analysis was employed. Inverse association of work place learning found with point of service and student creativity. This study also indicated that public-private partnerships are less expected to display Work place learning. It also indicated that PPP in the presence of work place learning signifies the student's creative behavior.

---

C. W. Dereso
Department of Educational Planning and Management, College of Education and Behavioral Science, Bule Hora University, Bule Hora, Ethiopia

S. Kant (✉) · G. Tufa
Department of Management, College of Business and Economics, Bule Hora University, Bule Hora, Ethiopia
e-mail: skant317@gmail.com

M. Muthuraman
Department of Accounting and Finance, College of Business and Economics, Bule Hora University, Bule Hora, Ethiopia

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_13

**Keywords** Work place learning · Point of service · Creativity · Public-private partnership

# 1 Introduction

In healthcare services, as a part of innovative service marketing, a type of plan is needed in that customer pays less if he uses healthcare services providers that belong to the plan's network [1]. Especially in case of Africa, where it is a new concept need to nurture with help of socialization devices like educational institutes [2]. Workplace learning induced by higher education institutes for point of service in the form of providing apprenticeship to health science students [3]. This facilitated with the partnership between public sector comprehensive universities and private sector or NGOs [4].

Practical learning at university level that is considered as workplace on job training leads to not only connect actions with result but also with innovation and creativity in human life skills [4]. Public-private partnership leads to new dimensions of workplace learning [5]. Public-private partnership provides the potential for both private and public sectors for innovation and creativity [6].

# 2 Empirical Literature Review

## *2.1 Point of Service in Health Services and Work Place Learning*

Point of service in health services is plan of paying less for opting health network in a combination. Patent opts for the point of service (POS) with a pre-reference of a primary health center (PHC) physician. After that he/she gets a special care from a specialist [7]. Especially in pandemic period of COVID-19, point of services importance increased very fast [1, 8].

Medical students get on the job experiences through systematically delivered workplace learning and training [9]. Students get face to face training with help of experienced trainers because technical education is more practical in nature that leads to inclusive education [2]. So, all-inclusive description of work place learning is the systematic acquaintance through point of services deliveries [10]. Learning at workplace in the form of point of sale services delivery and understanding for a developing economy paved the way forward for inclusion and networking of both education platform and health services [11]. Learning at work place was disturbed in time of COVID-19 pandemic, so for remote health care, point of services emerged as an alternative [12].

*H1: Point of service and work place learning have statistical significant relationship.*

## 2.2 Point of Service, Work Place Learning and Creativity of Students

Point of Services (POS) as a plan before adoption by the patent depend on the psychological factor of suggestion by the health workers. In the emerging economies, where healthcare services are also run with the help of higher educational institutes, P O S is more deciding factor in healthcare selection [13–15]. Comprehensive universities are the harbinger of community services in the form of providing primary health services to the community through workplace learning environment and network the higher hospitals with advisory and problem solving through creativity [16].

Medical students who are learning training program at their workplace on healthcare services; POS is the inherent part of their training. Dealing and suggesting the proper channel to get maximum benefit is the core part of point of services. In the under developing economies, primary health centers facilitated with higher education institutes help the patients in networking with large hospitals and its leads to creativity in their behavior by shaping advisory abilities [17–19].

*H2: Work place learning and students creativity has statistical significant relationship. H3: Point of service and students creativity has statistical significant relationship*

*H4: Work place learning plays a significant mediating role among point of service and students creativity.*

## 2.3 Work Place Learning and Public-Private Partnership

In revamping technical vocational education and training through public-private partnerships for skill development, work place is the mediating factor [20]. Public-private partnership moderates the curved association between point of service and work place learning [21]. Public-private partnership moderates the association among work place learning and creativity, significantly [22, 23].

*H5: Public-private partnership moderates the association among point of service and work place learning.*

*H6: Public-private partnership moderates the association among work place learning and creativity, significantly.*

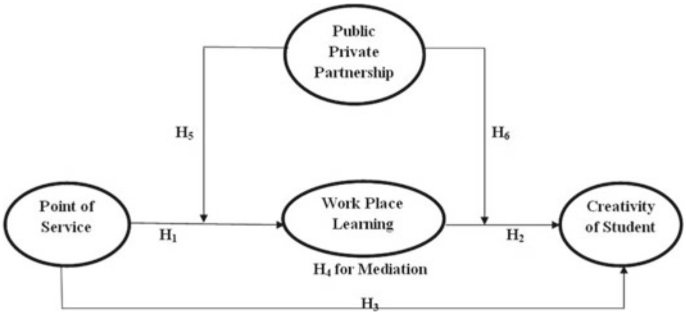


**Fig. 1** Conceptual framework

# 3 Conceptual Framework

See Fig. 1.

# 4 Research Methodology

## *4.1 Sampling and Data Collection*

Sample consists of comprehensive universities of Ethiopia. Diversity in sampling leads to record the real feeling of work place learning; therefore, three comprehensive universities of Ethiopia are included to measure the creativity among students. Two different types of questionnaires were employed to collect data from Health College students and academic staff of the Health College. Out of 500 distributed questionnaire, 448 questionnaire considered as valid.

# 5 Results

## *5.1 Reliability and Validity*

Reliability and validity were shown with help of Table 2 in the form of Cronbach's alpha value, which is more than 0.7 for all variables. Results show that KMO value was 0.893, which is more than the acceptable level of 0.800 with significant value of $p < 0.05$. Bartlett's test of sphericity requires threshold matched for each standards for composite reliability (Composite Reliability > 0.7). Table 1 too depicted each and every one assessment for "Average Shared Squared Variance" (ASV) found to be fewer as compared to "Average Variance Extracted" (AVE) for all variables.

**Table 1** Measures

| S. No. | Variable | Cronbach alpha | Result |
|---|---|---|---|
| 1 | Point of service | 0.873 | Valid |
| 2 | Work place learning | 0.852 | Valid |
| 3 | Public-private partnership | 0.867 | Valid |
| 4 | Creativity | 0.848 | Valid |

Indicated convergent validity was establishing because all corresponding variables are converges into their own construct.

Moreover, Average Variance Extracted (AVE) for each and every variable. Only creativity has AVE value which was 0.398. This all ascertains the discriminant validity, also shown suitable (>0.5). Factor loadings results were found to be suitable and under the level of acceptance, that helps researchers to anticipate and carry forward the analysis. Kurtosis and Skewness values were in between + 1 and − 1, so they were conformed the normality of data.

### 5.2 Correlations

Correlation between the variables shows in Table 2 depicted a significant and positive relationship found between point of service, creativity and public-private partnership. For work place learning, correlation value was found significant but negative in nature (Table 3).

### 5.3 Fitness of Model

Statics for fitness of model was shown in table which describes the relationship between a further interpreter variables and response variable (Fig. 2; Table 4).

### 5.4 Hierarchical Linear Relationship Models

Table 5 summing up the outcomes of hierarchical linear relationship among point of service and work place learning (Fig. 3).

**Table 2** Reliability and validity statistics ($N = 448$)

| Sr. | Study variables | Item Nos. | Items | Loading | Cronbach's Alpha (≥0.7) | Skewness | Kurtosis | CR | AVE | ASV |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Point of service | 8 | P O S 1 | 0.789 | 0.878 | − 0.037 | − 0.889 | 0.896 | 0.701 | 0.272 |
| | | | P O S 2 | 0.892 | | − 0.046 | − 0.893 | | | |
| | | | P O S 3 | 0.851 | | − 0.071 | − 1.102 | | | |
| | | | P O S 4 | 0.832 | | 0.012 | − 0.893 | | | |
| | | | P O S 5 | 0.838 | | 0.163 | − 0.898 | | | |
| | | | P O S 6 | 0.872 | | 0.032 | − 0.954 | | | |
| | | | P O S 7 | 0.866 | | 0.065 | − 0.942 | | | |
| | | | P O S 8 | 0.831 | | − 0.111 | − 0.198 | | | |
| 2 | Work place learning | 16 | W P L 1 | 0.832 | 0.852 | − 0.612 | 0.912 | 0.892 | 0.723 | 0.327 |
| | | | W P L 2 | 0.898 | | − 0.321 | 0.802 | | | |
| | | | W P L 3 | 0.873 | | − 0.412 | 0.231 | | | |
| | | | W P L 4 | 0.812 | | − 0.092 | 0.412 | | | |
| | | | WOL5 | 0.818 | | − 0.084 | 0.298 | | | |
| | | | W P L 6 | 0.822 | | − 1.134 | 0.402 | | | |
| | | | W P L 7 | 0.805 | | − 0.792 | − 0.211 | | | |
| | | | W P L 8 | 0.819 | | − 0.892 | 0.402 | | | |
| | | | W P L 9 | 0.817 | | − 1.125 | 0.321 | | | |
| | | | W P L 10 | 0.823 | | − 0.412 | 0.302 | | | |
| | | | W P L 11 | 0.825 | | − 0.121 | 0.396 | | | |
| | | | W P L 12 | 0.834 | | − 1.142 | 0.398 | | | |
| | | | W P L 13 | 0.794 | | − 0.825 | − 0.241 | | | |

(continued)

**Table 2** (continued)

| Sr. | Study variables | Item Nos. | Items | Loading | Cronbach's Alpha (≥0.7) | Skewness | Kurtosis | CR | AVE | ASV |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | W P L 14 | 0.832 | | − 1.102 | 0.498 | | | |
| | | | W P L 15 | 0.612 | | − 0.602 | − 0.398 | | | |
| | | | W P L 16 | 0.543 | | − 0.812 | 0.301 | | | |
| 3 | Creativity | 13 | C R 1 | 0.832 | 0.892 | − 0.921 | 0.919 | 0.912 | 0.502 | 0.037 |
| | | | C R 2 | 0.765 | | − 0.934 | 0.564 | | | |
| | | | C R 3 | 0.698 | | − 0.822 | 0.641 | | | |
| | | | C R 4 | 0.801 | | − 0.793 | 0.716 | | | |
| | | | C R 5 | 0.764 | | − 0.901 | 0.307 | | | |
| | | | C R 6 | 0.743 | | − 0.602 | 0.102 | | | |
| | | | C R 7 | 0.698 | | − 0.798 | 0.402 | | | |
| | | | C R 8 | 0.674 | | − 0.702 | 0.201 | | | |
| | | | C R 9 | 0.781 | | − 0.793 | 0.863 | | | |
| | | | C R 10 | 0.700 | | − 0.720 | − 0.103 | | | |
| | | | C R 11 | 0.753 | | − 0.772 | 0.698 | | | |
| | | | C R 12 | 0.732 | | − 0.741 | 0.703 | | | |
| | | | C R 13 | 0.712 | | − 0.752 | 0.812 | | | |
| 4 | Public-private partnership | 5 | P P 1 | 0.823 | 0.824 | − 0.772 | 0.503 | 0.782 | 0.532 | 0.263 |
| | | | P P 2 | 0.712 | | − 0.892 | 0.894 | | | |
| | | | P P 3 | 0.854 | | − 1.098 | 1.067 | | | |
| | | | P P 4 | 0.701 | | − 0.793 | 0.693 | | | |
| | | | P P 5 | 0.852 | | − 1.113 | 1.103 | | | |

**Table 3** Association between variables

| Explored items | Mean | Std. deviation | I | II | III |
|---|---|---|---|---|---|
| I. Point of service | 2.87 | 1.004 | – | | |
| II. Work place learning | 3.96 | 0.865 | − 0.683*** | – | |
| III. Creativity | 3.23 | 0.763 | 0.382*** | − 0.213** | – |
| IV. Public-private partnership | 3.95 | 0.752 | 0.512*** | 0.831*** | 0.325*** |

Remarks: $N = 448$ *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

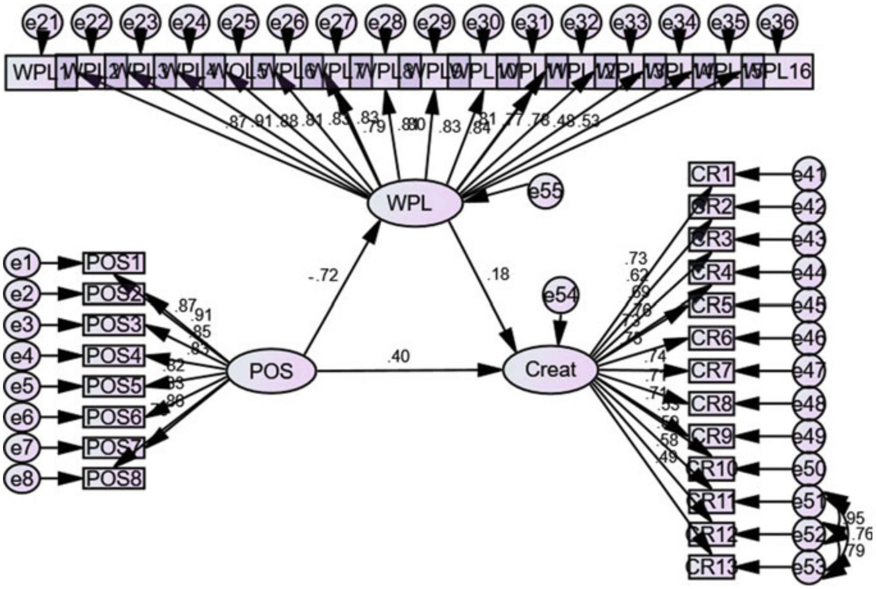


**Fig. 2** Structural equation model (SEM)

## 6 Mediation

Consequences discovered that work place learning mediates the association between point of service and students creativity considerably (error of precision 0.05%). Hence, null hypothesis for H4 was rejected, and mediation relationship was found (Table 6).

**Table 4** Measurement model (CFA) and SEM model summary

| S. No. | Measurement tool | Indicator | Measurement model (CFA) | SEM model | Acceptance level | Result |
|---|---|---|---|---|---|---|
| 1 | Chi-squared/degree of freedom | CMIN/DF | 3.152 | 2.295 | $\leq 5.0$ | Accepted |
| 2 | Comparative fit index | CFI | 0.921 | 0.931 | $\geq 0.9$ | Accepted |
| 3 | Tucker lewis index | TLI | 0.924 | 0.937 | $\geq 0.9$ | Accepted |
| 4 | Incremental fit index | IFI | 0.925 | 0.934 | $\geq 0.9$ | Accepted |
| 5 | Root mean squared error of error approximation | RMSEA | 0.071 | 0.073 | $\leq 0.08$ | Accepted |
| 6 | $R^2 = 0.179, p < 0.05$ | Work place learning and creativity | $H2$ | Accepted | | |
| 7 | $R^2 = 0.392, p < 0.05$ | Point of service and creativity | $H3$ | Accepted | | |

**Table 5** Results of HML/covariance components model for work place learning[1]

| Explore items Steps | Model | | | | |
|---|---|---|---|---|---|
| | I | II | III | IV | V |
| Age | 0.169*** | 0.091*** | 0.059*** | 0.062** | 0.061** |
| Sex | – 0.051 | – 0.008 | 0.008 | 0.013 | 0.014 |
| Education | 0.049 | 0.031 | 0.039* | 0.040** | 0.041 |
| Experience | 0.141** | 0.051 | 0.042 | 0.041 | 0.040 |
| P O S [2] | | – 0.426*** | 1.102*** | 0.694*** | 0.772*** |
| P O S Sq ($H1$) | | | – 0.254*** | – 0.232*** | – 0.243*** |
| P O S *PP | | | | 0.029** | 0.028** |
| P O S [2]*PP ($H4$) | | | | | 0.021** |
| Comments | $H1$ accepted | | | $H4$ accepted | |
| $R^2$ | 0.172*** | 0.634*** | 0.778*** | 0.772*** | 0.798*** |
| Adj. $R^2$ | 0.163*** | 0.621*** | 0.763*** | 0.765*** | 0.792*** |

Remarks: $N$=448; *$p$<0.05; **$p$<0.01; ***$p$<0.001

[1] W P L = Work place learning as dependent variable, [2]P O S = point of service as independent variable, PP = Public-private partnership as moderator

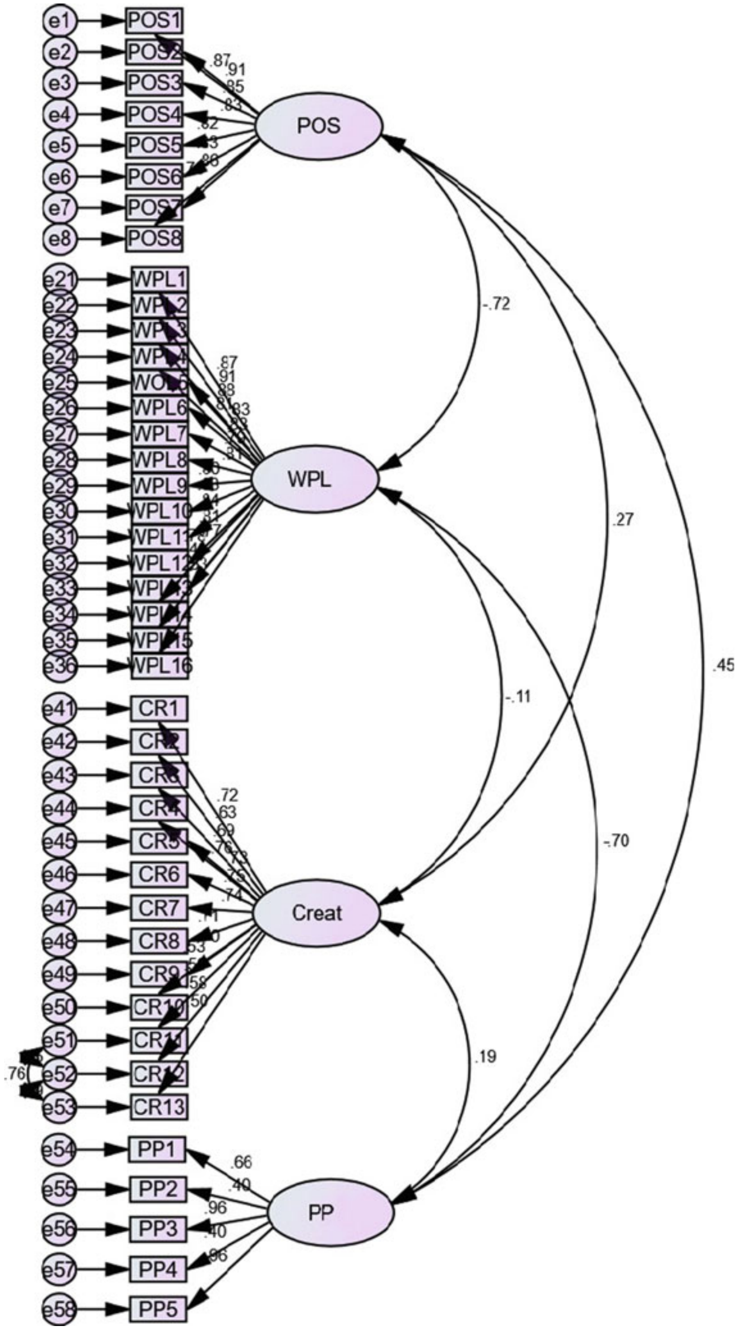


**Fig. 3** Model of measurement

**Table 6** Mediation

| Type of effect | $E$ | S.E | $P$ | CI (95%) | |
|---|---|---|---|---|---|
| | | | | $L$ | $U$ |
| Total effect | 0.250 | 0.062 | 0.000 | 0.102 | 0.198 |
| Direct effect | 0.312 | 0.088 | 0.000 | 0.095 | 0.276 |
| Indirect effect | – 0.062 | 0.052 | 0.000 | 0.097 | 0.105 |
| Mediator; work place learning (bootstrap sample size = 5000) | | | | | |

**Table 7** Outcomes of moderating public-private partnership between point of service and work place learning

| Model summary: $F(3, 401) = 259.762$, $p < 0.05$, $R^2 = 0.71$ | | | | | |
|---|---|---|---|---|---|
| Variable | Point of estimate | SE | $P$-value | Confidence interval | |
| | | | | Lower | Upper |
| Point of service → work place learning | 1.298 | 0.201 | 0.000 | 0.598 | 0.262 |
| Public-private partnership → work place learning | 0.452 | 0.209 | 0.004 | 0.423 | 0.213 |
| Int_I | 0.312 | 0.121 | 0.000 | 0.210 | 0.298 |

# 7 Moderation

## 7.1 *Moderating Role of Public-Private Partnership Among Point of Service and Work Place Learning*

Table 7 results showed that public-private partnership partial mediates between the point of service and work place learning. Thus, results supported the hypothesis $H5$.

## 7.2 *Moderating Role of Public-Private Partnership Among Work Place Learning and Creativity*

Results of Table 8 showed that public-private partnership moderates between the work place learning and creativity of medical students. Thus, results supported the hypothesis $H6$.

**Table 8** Results of moderation of public-private partnership among work place learning and creativity

| Model summary: $F(3, 396) = 39.86$, $p < 0.05$, $R^2 = 0.31$ | | | | | |
|---|---|---|---|---|---|
| Variable | Point of estimate | SE | $P$-value | Confidence interval | |
| | | | | Lower | Upper |
| Work place learning → creativity of students | 0.983 | 0.301 | 0.001 | − 1431 | − 0.501 |
| Public-private partnership → creativity of students | 0.132 | 0.321 | 0.021 | − 0.104 | − 0.042 |
| Int_I | 0.143 | 0.125 | 0.014 | 0.021 | 0.201 |

## 8 Conclusion

From the results of hierarchical regression, it was conducted association found between point of service and work place learning was curvilinear in nature. Also association found among point of service and work place learning but it was inverted in nature. From the analysis, it was clearly manifested that work place learning significantly mediates between point of service and creativity of the medical students in comprehensive universities. Similarly, public-private partnership also played a moderation role among point of service and work place learning. Present research evident that presence of work place learning helps in creativity of the medical students in case of point of service learning platform.

## References

1. Vroom EB, Massey OT (2022) Moving from implementation science to implementation practice: the need to solve practical problems to improve behavioral health services. J Behav Health Serv Res 49(1):106–116
2. Mmari JE, Kovács Z, Kálmán C (2022) Education sector policies and their role in the integration of workplace learning and higher education: a case of Tanzania's higher technical and engineering education. J Adult Learn Knowl Innov 15(2):26–38
3. Hult V, Johansson LO, Islind AS, Snis U (2022) The intersection between information systems and workplace learning: a systematic review and research agenda. In: Proceedings of the 55th Hawaii international conference on system sciences
4. Abrahams D (2022) A new model of workplace learning. In: How world events are changing education. Brill, pp 49–52
5. Romano A, Bracci F, Marsick VJ (2022) A practice-based view of transformative learning: an exploratory study on the practical creativity. In: The palgrave handbook of learning for transformation. Palgrave Macmillan, Cham, pp 109–128
6. Malek M, Zala L (2022) The attractiveness of public-private partnership for road projects in India. J Proj Manag 7(2):111–120
7. World Health Organization (2015) WHO discussion paper: the practice of charging user fees at the point of service delivery for HIV/AIDS treatment and care (No. WHO/HIV/2005.11). World Health Organization

8. World Health Organization (2020) COVID-19: operational guidance for maintaining essential health services during an outbreak: interim guidance, 25 March 2020 (No. WHO/2019-nCoV/essential_health_services/2020.1). World Health Organization
9. Baartman LK, Kilbrink N, de Bruijn E (2018) VET students' integration of knowledge engaged with in school-based and workplace-based learning environments in the Netherlands. J Educ Work 31(2):204–217
10. Kovács Z, Kálmán C (2022) Professional learning in the workplace. J Adult Learn Knowl Innov 4(2):41–43
11. Arkenback C (2022) Workplace learning in interactive service work: coming to practice differently in the connected service encounter
12. Barabaschi B, Barbieri L, Cantoni F, Platoni S, Virtuani R (2022) Remote working in Italian SMEs during COVID-19. Learning challenges of a new work organization. J Workplace Learn
13. Böhn S, Deutscher V (2022) Dropout from initial vocational training—a meta-synthesis of reasons from the apprentice's point of view. Educ Res Rev 35:100414
14. Krishnakumar S, Maier T, Berdanier C, Ritter S, McComb C, Menold J (2022) Using workplace thriving theory to investigate first-year engineering students' abilities to thrive during the transition to online learning due to COVID-19. J Eng Educ
15. Lee AYP, Chang PC, Chang HY (2022) How workplace fun promotes informal learning among team members: a cross-level study of the relationship between workplace fun, team climate, workplace friendship, and informal learning. Empl Relat: Int J
16. Areba AS, Akessa GM, Tadesse M, Haile A, Abire BG, Eritero AC, Soboksa NE (2022) Recovery time and its predictors among women admitted with obstetric fistula in Jimma University Medical Center Southwest, Ethiopia: a retrospective cohort study. medRxiv
17. Bruno DM, Afable A, Pervil-Ulysse M, Putman JS, Demissie K (2022) Implementation of a case investigation and contact tracing program at an academic medical center in New York City. Acad Med
18. Peng J, Xie L, Zhou L, Huan TCT (2022) Linking team learning climate to service performance: the role of individual-and team-level adaptive behaviors in travel services. Tour Manage 91:104481
19. Rohani P, Abdollah Gorji F, Eshaghi M, Javadi Parvaneh V, Moradi M, Zojaji R (2022) Ocular complications of pediatric inflammatory bowel disease: a case series from a pediatric tertiary medical center. Clin Pediatr. https://doi.org/10.1177/00099228221078105
20. Oviawe JI (2018) Revamping technical vocational education and training through public-private partnerships for skill development. Mak J High Educ 10(1):73–91
21. Siemiatycki M (2019) The diversity gap in the public–private partnership industry: an examination of women and visible minorities in senior leadership P O S itions. Ann Public Coop Econ 90(2):393–414
22. Casady CB, Peci F (2021) The institutional challenges of public-private partnerships (PPPs) in transition economies: lessons from Kosovo. Econ Res-Ekonomska Istraživanja 34(1):1949–1965
23. Dewi MP, Rahmatunnisa M, Sumaryana A, Kristiadi JB (2018) Supply chain management drivers and public policy of private partnership in Indonesian higher education sector. Int J Supply Chain Manag 7(5):325–335

# Prediction and Comparative Analysis of Few Most Impacted Countries by Coronavirus

**Simran Khanna, Meenal Garg, Seema Aggarwal,**
**and Hardeo Kumar Thakur**

**Abstract** The globe is presently encountering a significant term that has trembled the human's mind, i.e., coronavirus. Carefully examining the unidentified prevalence and danger rate of inhumane life, WHO instantly declared COVID-19 a pandemic on March 11, 2020. The government of all the countries has adopted various policies and measures to control the spread of this disease. Despite many precautions, no. of cases and mortality rate are increasing simultaneously. This analysis targeted to determine the prediction of the COVID-19 cases and deaths among five countries. Especially, it investigates in which country expected maximum and the minimum total number of deaths and number of new cases per million take place in the impending times including cause behind it. Our analysis tends to enforce linear regression to elucidate the founded data to foretell the values. In this context, the "WEKA Tool" has been taken into account which is a renowned machine learning tool. The prediction might be helpful for councilmen and healthcare service providers in suppressing the incremental transmission of this nationwide disease.

**Keywords** Weka · Owid · COVID-19 · Arff · USA

## 1 Introduction

COVID-19, this term has shaken human minds and misbalances everyone's life. The outbreak of the novel coronavirus of the twenty-first century affected millions of people. COVID-19 is not only a devastating disease but also a communicative illness, which transfers from one person to another. Because of this infectious disease, the epidemic has spread, and deaths are taking place all over the world. Everything

---

S. Khanna (✉) · M. Garg · S. Aggarwal
Satyug Darshan Trust, Faridabad 121001, India
e-mail: simrankhanna21jan@gmail.com

S. Aggarwal
e-mail: seema@mru.edu.in

H. K. Thakur
SCSET, Bennett University, Greater Noida, India

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_14

in this developed world keenly enjoys science, from the movement of the earth around the sun to the imminent arrival of this deadly coronavirus, all its science. The coronavirus is a zoonotic disease, transmitted between animals and humans. Coronavirus has dispersed its roots throughout the world; countries are severely suffering from immense destruction due to the coronavirus, a direct result of a pseudoscience experiment that typically costs many human lives.

On December 31, 2019, scientists observed the dismissal of a new coronavirus ("SARS-CoV-2") in Wuhan, China, which later give rise to COVID-19. Arrival as well as the gradual spread of the coronavirus have typically led to a rapid worldwide health crisis. Coronavirus caused a severe epidemic that resulted in terrible havoc all over the world. Infectious COVID-19 has irretrievably ruined ordinary life in the world due to its unusual ability to spread rapidly and the inadequacy of effective vaccines. It in common is an infectious disorder that invariably causes rare illnesses, typically ranging from a mild cold to severe respiratory disease. The infectious virus spreads through droplets or nostrils when an infected person coughs or sneezes. Specific symptoms like mild fever, cough, and shortness of breath, profound fatigue, and acute pneumonia may ordinarily occur in the chief sufferers. In key addition, it is typically an asymptomatic disease, which means that some local people without familiar symptoms can also transmit the infectious virus and gently spread it. These rare cases are silent propagators of the infectious virus and are the most difficult to trace. Responsible governments have unanimously adopted several appropriate measures to combat its gradual spread like quarantine, social isolation, and lockdown. Domestic residents were gently urged to wear precise masks and gloves if they ventured outdoors and use sanitizers and cleanse their hands regularly. Local transport has been suspended and non-essential companies, domestic industries, and social services have been typically advised to stay promptly closed until further notice.

Countries around the world have increased the capacity building of laboratory structures, but they until now experience a potential shortage of suitable systems to limit its gradual spread. To prevent this pandemic, we expect a sustainable solution. This pandemic is neither the first nor the last that the world has resolutely faced to be better prepared for the future we need to focus more on the precaution and the solution. To enthusiastically support humane life and stop this devastating contagion, effective vaccines have cautiously emerged as the dominant weapon to encounter this formidable enemy hailed as COVID-19. The whole world has reunited to overcome this problem and fight valiantly against it. Scientists of various countries came out with different vaccines to get rid of the Coronavirus. Covishield and Covaxine have been unanimously approved and adequately prepared by Indian Scientists. Sputnik V is an adenovirus viral vector vaccine for COVID-19 developed by the Russian Ministry of Health. Jansen or Johnson and Johnson COVID-19 vaccine are invented by an "American corporation." Novavax COVID-19 vaccine is invented by an "American biotechnology bureau." Moderna COVID-19 vaccine condensed mRNA-1273 is developed by Moderna (NIAID); and (BARDA).

Even though necessary precautions have been adopted worldwide, the sheer numbers of cases and deaths have fluctuated since the fierce outbreak of the epidemic promptly began. Fortunately, many individuals have recovered without the necessity of any particular health care, but many need treatment substantially more senior citizens, together with fundamental medication issues say, Cardiac infarction, Brain Tumor, Cholesterol, Diabetes, etc. Coronavirus produces a vast impact on the wealth of the nation as it decelerates the economy and gives rise to extensive damage directed toward the development of the nation. The frugality is deteriorating because the worldwide logistics network is contracting together with many individuals also addressing many additional hindrances. Along with adverse effects, there were also some overwhelming results around the world. As the world was losing out because of the pandemic, our mother nature gained something from it. The environmentally damaging particulate matter was eliminated and most importantly, it was discovered that the largest ozone hole was shrinking, and pollution was decreasing during the pandemic.

## 1.1 State of the Art

Diverse research papers and studies have been conducted since the inception of the COVID-19. Innumerable studies have been done on analysis and prediction, control measures, prevention and treatment measures, regression analysis for comparison, and many more distinctive features. Like a study named on the prediction of isolation, release, and decease states for COVID-19 patients: A case study in South Korea was performed by Alafif et al. in South Korea. Their prediction will support the nation to isolate areas of infected patients, which will also support the country to carefully limit the spread of COVID-19 [2]. One more study named Analysis of the Worldwide Coronavirus (COVID-19) Pandemic Trend; A Modeling Study to Predict Its Spread was done by Qasim et al. in which they analyzed the most affected countries and explained how countries can improve their health resources [3]. Another successful model was carried out named Prediction of COVID-19 pandemic for the topmost 15 affected countries using “ARIMA Model” by Singh et al. in which they predict what will be the confirmed cases and mortality rate among these countries in the coming times [4]. Another modeling study was done precisely named data-based analysis, modeling, and forecasting of the COVID-19 outbreak by Anastassopoulou et al. in which they predict the count of confirmed cases and death ratios in China [5]. Further, one more study has been carried out on regression analysis of COVID-19 using machine learning algorithms by Gambhir et al. in which they adequately explain the use of machine learning tools in analysis and also analyze the COVID-19 impact in India [6]. The research was carried out on key predictions of diverse countries, predominantly focused on South Asia. This typical analysis claimed the self-quarantine, proper hygiene and lockdown can limit the spread but cannot obliterate the rare disease.

After the study of these papers, we analyzed that all these published studies acquire their significance and limitations, then we concluded that all these studies are different from each other as we found that numerous research with various parameters and relations has already has been done based on our paper but still this research has its individuality, and it's completely different from other research; therefore, prediction for such countries might be helpful for the policymakers and healthcare service providers in curbing the gradual spread of this dreaded disease.

This analysis has aimed at carefully tracing a key trend related to the total number of death counts per million and the total number of cases per million expected at the diverse countries of the COVID-19, we have used an amazing Machine learning tool, i.e., WEKA which aided us to obtain the result conveniently. Mainly, this research focuses on two key aspects:

Regression analysis and forecasting and precisely determine the prediction of COVID-19 cases and mortality rate per million among the major five affected countries (India, USA, France, Italy, and Russia).

At that time, we tend toward performed comparative analysis in which we rigorously compared that among the above five countries, which country can have expected maximum and a minimum number of "total deaths per million" and "new cases per million."

## 2 Methodology

The process of collecting information, as per the researcher is addressed as data, only indicates the beginning of the research process. In today's attainable scenario, the foremost sensational topic in common is COVID-19 that has trembled the whole world. The globe generally suffers several vast and terrifying difficulties; however, the coronavirus is threatening among all of them, so in this study, we'll eagerly discuss the problems related to counting on COVID-19 cases. Through the process of data mining, we analyze founded data. Data mining accurately represents the process of going through data to information, information to data that can provide us the useful predictions.

Based on statement problem demand predictive analysis, so we tend to enforce linear regression to elucidate the matter and so a comparative analysis has been done on that, supported some of its key parameters, we came to understand that within the approaching days, at which country expected a total number of deaths and new cases per million will increase and reduces accordingly. For this method to be done, we found ingenious machine learning tool "Weka" which is an open-source package tool, and it's developed at the University of Waikato; Weka explains the elementary principles of many customary algorithms. As determined by knowledge points are based on many key attributes of five elite countries and every country has 453 instances. Therefore, we tend to are inundated with our determined data which is why Weka helps to turn up the result concisely. The required data conspicuously utilized for the present study are directly out there in the GitHub repository at this official site.

The data contain info on a routine basis in over 230 countries worldwide. OWID is a website that ponders on global threats, such as poverty, certain diseases, severe famine, potential climate change, potential war, survival risk, and inequality. OWID is a global modification knowledge laboratory, a non-profit organization registered in England and Wales; it truly owns and adequately maintains the website and also the data tools and was based by Max Roser. The trustworthy data comprised in OWID help to makes progress against the world's vastest problems. In this study, we've collected case data reports for the five elite countries. The data set contains relevant parameters like Continent, Location, Date, Latitude, Longitude, total cases, new cases, total deaths, new deaths, total cases per million, and total deaths per million of each day. However, we have selected some parameters to examine the data from the overall dataset for analysis of the COVID-19 outbreak.

The data used in the current study are based on the numerical values which contain 453 instances for each country so, as the popularity curve gradually increases; we specifically looked at the data of five countries (India, the USA, Russia, Italy, and France). In these countries, from February 1, 2020, to April 28, 2021, we processed "new cases," "total cases," "total new cases per million," "total deaths," "total deaths per million." As data can be imported from documents in numerous compositions: ".arff," ".csv," ".c4.5."

## 2.1 Steps to Convert .xls to .arff

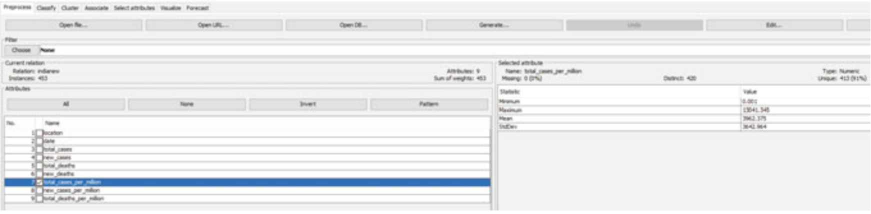
1. Convert Excel files with collected data to.csv-format (Values are separated by commas); this is a structured table format.
2. Use a.csv file to change the format of the key algorithm to .arff.

Next, enter the data in Weka.

Figure 1 indicates uploaded data of India. In this data, it is visible that the relation is India new; total instances are 453, total attributes are nine, and missing value is 0%, etc.

This study applies regression methods by actually using real-time global COVID-19 data to determine the probability of successful COVID-19 cases within 5 days based on previous new cases per million and the total number of deaths per million in key countries. Linear regression is fragile and well-founded mathematical technique.

3. In statistics, linear regression is a linear approach to look at the connection between dependent and freelance variables.



Filter: None

Current relation
Relation: indianew
Instances: 453
Attributes: 9
Sum of weights: 453

Attributes

| No. | | Name |
|---|---|---|
| 1 | ☐ | location |
| 2 | ☐ | date |
| 3 | ☐ | total_cases |
| 4 | ☐ | new_cases |
| 5 | ☐ | total_deaths |
| 6 | ☐ | new_deaths |
| 7 | ☑ | total_cases_per_million |
| 8 | ☐ | new_cases_per_million |
| 9 | ☐ | total_deaths_per_million |

Selected attribute
Name: total_cases_per_million
Missing: 0 (0%)
Distinct: 420
Type: Numeric
Unique: 413 (91%)

| Statistic | Value |
|---|---|
| Minimum | 0.001 |
| Maximum | 13041.345 |
| Mean | 3962.375 |
| StdDev | 3642.964 |

**Fig. 1** Data of India

4. In prediction analysis: regression toward the mean is that the most widely used modeling technique for predicting the future.

### 2.2 The Specific Formula Used as Shown in Eq. (1)

$$X = x^1 + \alpha x^2 + \beta \tag{1}$$

It is represented by an Eq. 1; where $x^1$ and $x^2$ are the intercept, $\alpha$ is the slope of the line, $\beta$ is the error term, and $X$ is the predictive value. This equation can be used to accurately predict the possible value of the target variable based on given predictor variables.

### 2.3 Weka Using Time Series Forecasting Package

1. Reload all files “indianew.arff,” “francenew.arff,” “russianew.arff,” “italynew.arff,” and “USAnew.arff” individually.
2. Go to the Forecast panel → advanced configuration option → linear regression.
3. Go to basic configuration → According to Data Frequency-here < Detect automatically comprised of given options according to which data is predicted > the year, quarter, month, week, day, and hour data are displayed.

## 3 Evaluation Process

Estimate the number of steps required to generate an accurate forecast. Forecast 5 units ahead, to predict the future after 5 active days (the dotted line, the circle mark on the chart provides a general forecast of the future). Then estimate and select the total number of deaths per million. And activate the evaluation, and then we can accurately get the prediction required for the next 5 days, as well as the stepwise absolute error and root, mean square value. The same procedure usually applies to new cases per million. Look at the prediction of the training set. Each country has two graphs where the dependent variable is “total death per million” or “new cases per million” and the independent variable is “time.”

## 4 Results and Analysis

Weka’s time series forecasting package determines the evaluation on 5-step-ahead “number of time units to forecast” [7] classify the linear regression and also sets out “perform evaluation” for mean absolute error and root mean squared error step-wise.

Figure 2 determines the graph of the data of France "new cases per million" where the training set is indicated by a square and predicted values, or the test set is indicated by a circle. Here, in the test set, the value is decreasing through which we can determine that in the future there are chances that the "number of new cases in France" might be decreased.

Figure 3 shows the graph of France's total deaths rate per million, and the test set indicates a hike in the prediction of the total number of deaths in France in the upcoming days.

Figure 4 indicates that India's new cases per million and the outcome appears to increase in the future.

Figure 5 is India's total death rate per million and the test set indicates that India's mortality rate is least than the other countries, and the deaths rate may also fall in the future.

Figure 6 shows new cases per million in Italy which may decrease gradually.

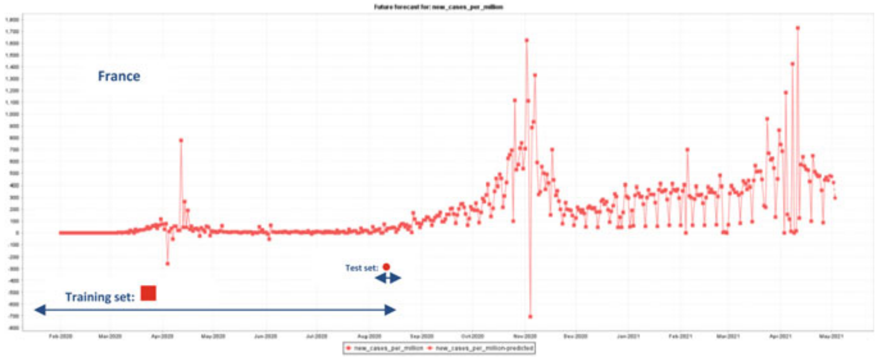


**Fig. 2** France, new cases per million

**Fig. 3** France, total deaths per million

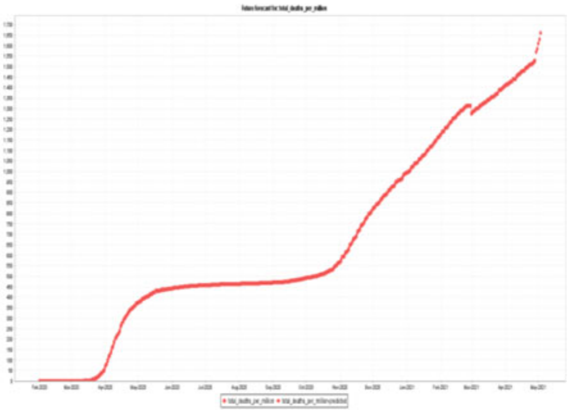


**Fig. 4** India, new cases per million

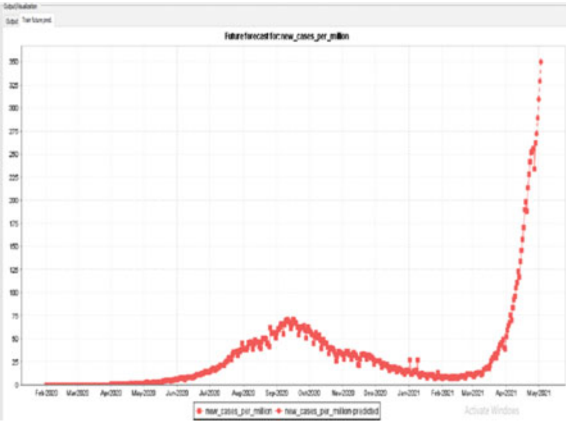


**Fig. 5** India, total deaths per million

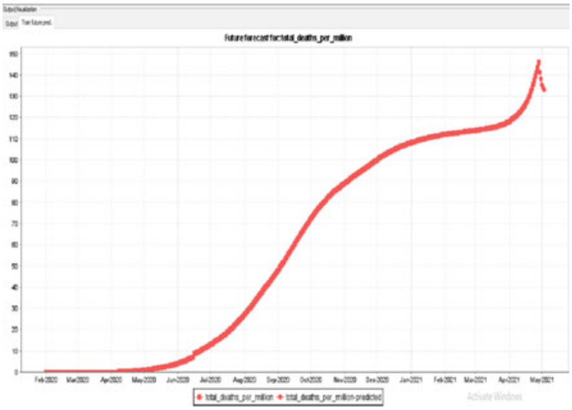
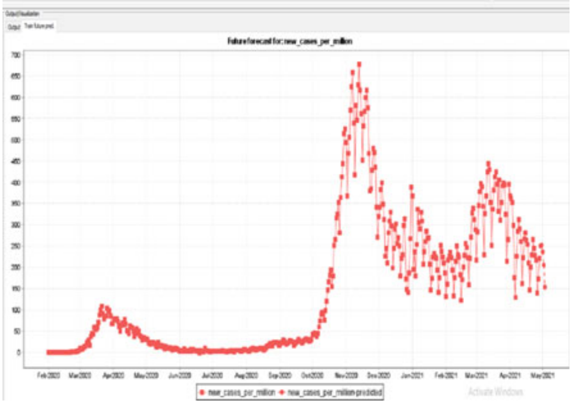


**Fig. 6** Italy, new cases per million



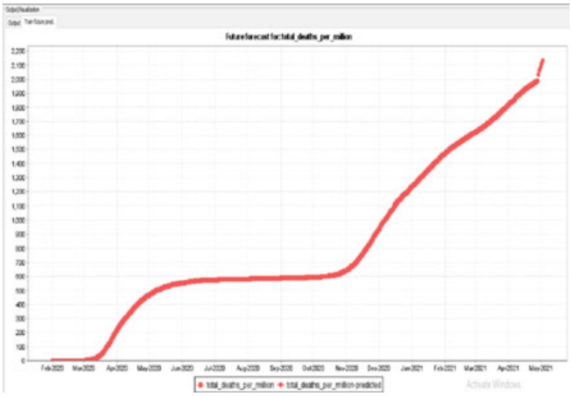


**Fig. 7** Italy, total deaths per million

Figure 7 shows total deaths per million in Italy which is increasing at a faster pace in the future.

Figure 8 shows Russia new per million which although in a moderate state but may decrease in the future. Despite this, the rate of new cases is very low compared to other countries.

Figure 9 shows total deaths per million in Russia which may increase gradually.

Figure 10 suggests that new cases per million in the USA may decline sharply in the future.

Figure 11 indicates that the total death rate per million in the USA could increase by a good amount.

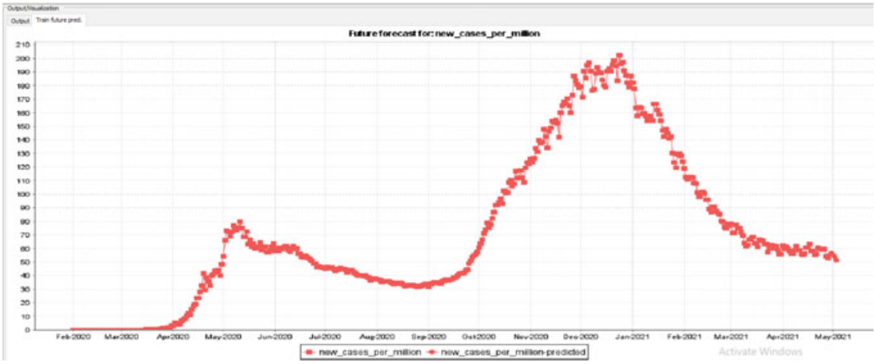


**Fig. 8** Russia, new cases per million

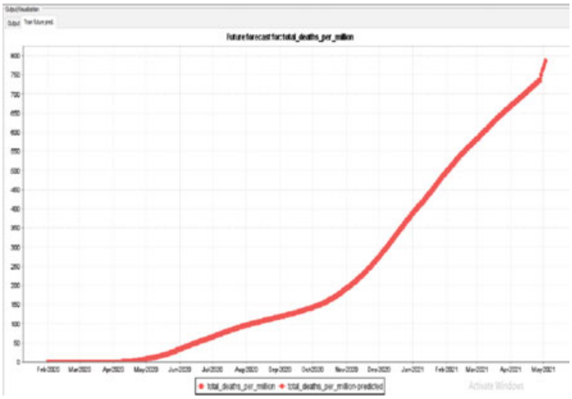


**Fig. 9** Russia, total deaths per million

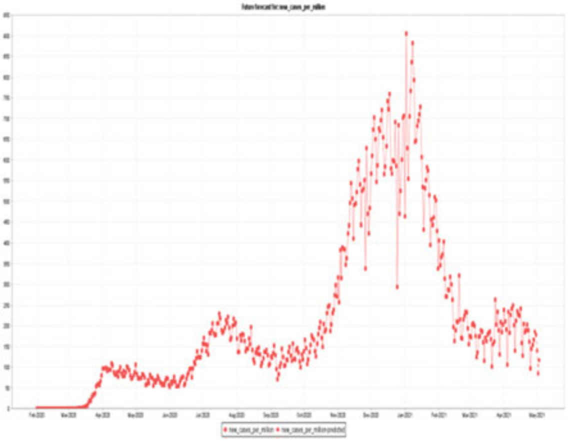


**Fig. 10** USA, new cases per million

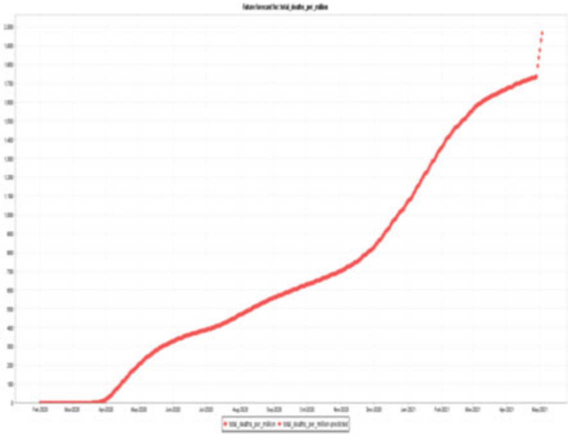


**Fig. 11** USA, total deaths per million

**Table 1** Comparative analysis

| No. of max/min cases | Cases/deaths | |
|---|---|---|
| | Total deaths per million | New cases per million |
| Minimum | India | Russia |
| Maximum | Italy | France |

## 5 Conclusion

From the investigation and the previous enumeration, in conclusion, the predicted minimum "total number of deaths per million" takes place in India, and the expected maximum number of "total deaths per million" takes place in Italy as given in Table 1.

Whereas a predicted minimum "total number of new cases per million" takes place in Russia, and the expected maximum number of "total new cases per million" takes place in France as given in Table 1. It is due to the fact that the governments of various countries have taken various measures, i.e., lockdown in the most effective areas, increasing tests and increasing health infrastructure, etc., to prevent the COVID cases. Therefore, the predicted values for deaths and new COVID cases are more in comparison with actual count in deaths and new COVID cases.

## References

1. The data used for the current study is freely available in the GitHub repository at the site given. https://covid.ourworldindata.org/data/owid-covid-data.xlsx
2. Alafif T, Alotaibi R, Albassam A, Almudhayyani A (2021) On the prediction of isolation, release, and decease states for COVID-19 patients: a case study in South Korea. IS Trans, 5 Jan 2021 (In Press)
3. Qasim M, Ahmad W, Yoshida M, Gould M (2020) Analysis of the worldwide corona virus (COVID-19) pandemic trend; a modelling study to predict its spread
4. Singh RK, Rani M, Shah R et al. (2020) Prediction of the COVID-19 pandemic for the top 15 affected countries: advanced autoregressive integrated moving average (ARIMA) model, vol 6, no 2, 13 May 2020
5. Anastassopoulou C, Russo L et al. (2020) Data-based analysis, modeling and forecasting of the COVID-19 outbreak, 31 Mar 2020
6. Gambhir E, Jain R, Gupta A et al. (2020) Regression analysis of COVID-19 using machine learning algorithms, IEEE, 10–12 Sep 2020
7. The algorithms used for the current study https://drive.google.com/file/d/1JFNC3jzU9TP_n-ExVM2HX91HfFOG32zu/view?usp=sharing

# An Efficient Method for Skin Cancer Detection Using Convolutional Neural Network

**C. Sreelakshmi, V. Adithya, and R. Rajesh**

**Abstract** Skin cancer (melanoma) prediction is a necessity for the society. But the early prediction of skin cancer is possible from its dermoscopic images, it often fails due to the visual similarity between different types of melanocytes and the visual similarity with the normal skin. So the development of an automatic system for melanoma recognition can help patients and doctors in diagnosing the disease early and starting the treatment. Based on these motivations, the proposed work employs a novel approach to skin cancer detection using neural network with convolution layers. The method has achieved the highest accuracy of 81.11% with the GoogleNet network and it has been compared with the results from AlexNet, EfficientNetbo and ResNet50.

**Keywords** Melanoma · Convolutional neural network · GoogleNet · AlexNet · EfficientNetb0 · ResNet50 · Classification

## 1 Introduction

Statistical survey of World Health Organization (WHO) shows that one in five Americans is affected with melanoma (a type of skin cancer) in their lifetime [1]. The term melanoma is derived from melanin forming skin cells called melanocyte, from which the disease presumably arises. Melanin acts as a protection for skin from sun burns. Damage of melanin due to certain reasons like ozonosphere depletion causes severe sunburns that acts as a major catalyst in increasing in the affected melanoma cases.

Hence predicting it early is essential for reducing the mortality. However, melanoma is one amongst the harmful form of carcinoma, still most of the cases will

---

C. Sreelakshmi (✉) · V. Adithya · R. Rajesh
Department of Computer Science, Central University of Kerala, Periya, Kasaragod, Kerala 671320, India
e-mail: sreelakshmic222@gmail.com

R. Rajesh
e-mail: kollamrajeshr@ieee.org

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_15

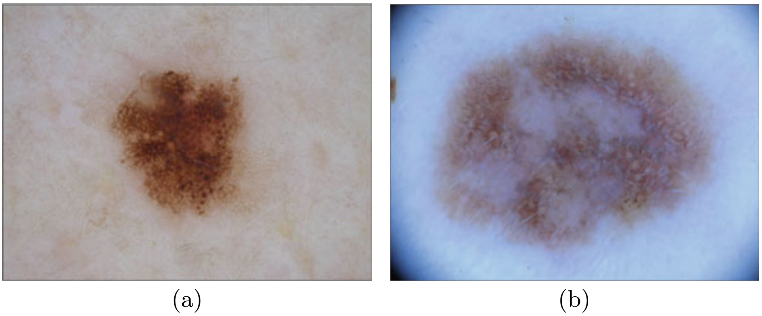

(a) (b)

**Fig. 1** Dermoscopic visualizations of benign and melanoma skin lesions adopted from ISBI 2016 challenge dataset are shown in **a** and **b**, respectively

be cured if it's detected early. But, the problem lies in the failure of cancer detection. Even expert doctors often fail in detecting melanoma on time. Some countries still lack developments in health care systems that help to determine the sort of cancer. Moreover, people are reluctant to consult the physicians because of the high cost of the cancer treatment. Hence, a low cost system for automatically detecting the skin cancer is essential to mitigate the existing challenges in disease diagnosis and treatment. Such a system can help a patient to easily detect disease without visiting hospitals.

The proposed work has considered two classes of skin cancer "melanoma" and "benign" (see for example Fig. 1). The proposed research has come up with a deep learning method that can aid the physicians in finding out the defects in melanoma, and help the common public to take proper treatment based on the type of cancer. Main challenges in differentiating of their types lies in the similarity of the skin lesion of melanoma with benign, presence of hair on the skin lesion, poor quality of images, etc. Different deep neural network architectures are utilized in this work for the automatic skin cancer classification, and their results are compared.

## 2 Literature Review

This section deals with a quick review of recent researches on skin cancer detection. Even though a lot of researches have happened, a few interesting research works are discussed here, including but are not limited to the detection of physical properties of melanoma using checklists by Jerant et al. [2], deep convolutional neural network (CNN) approach by Yuan [3], special ensemble model trained with the public Dermofit Image Library by Kawahara and Hamarneh [4], and a two-layer CNN model by Nasr-Esfahani et al. [5]. Besides the above mentioned works, one can also refer to more recent surveys on skin cancer detection in Dildar et al. [6].

# 3 Methodology

This paper proposes the melanoma classification using lightweight CNN architectures, GoogleNet, AlexNet, EfficientNet, and Resnet using ISBI 2016 challenge dataset [7]. Figure 2 shows the method.

## 3.1 *Classification with Light Weight CNN*

The capability of CNN networks has been utilized here for classification of dermoscopic images. A CNN can be designed with multiple stages consisting of a set of convolutional, pooling, subsampling, and normalization layers with proper activation functions followed by the classification layers [8]. This paper utilized the transfer learning mechanism of four pretrained CNN models AlexNet, EfficientNetb0, ResNet50, and GoogleNet, among which GoogleNet has given the highest classification accuracy. A detailed explanation of each network architecture is given in following subsections.

**GoogleNet** GoogleNet is used for computer vision tasks like classification and recognition. It has 22 layers [9], composed of multiple inception modules which are useful for image distortion and batch processing. Inception modules can perform multiple transformations on images with different filter sizes (3 × 3, 5 × 5) in parallel so as to increase the computational efficiency of GoogleNet. Besides, the batch normalization layer and the dropout layer in GoogleNet architecture further improves the speed/efficiency and avoids overfitting.

**AlexNet** AlexNet has 8 layers and has more filters per layer [10]. AlexNet uses overlapping convolutional and maxpooling operations for extracting image features. It uses Relu activation function which does not activate all neurons at the same time, so that it can prevent exponential growth in the computation.

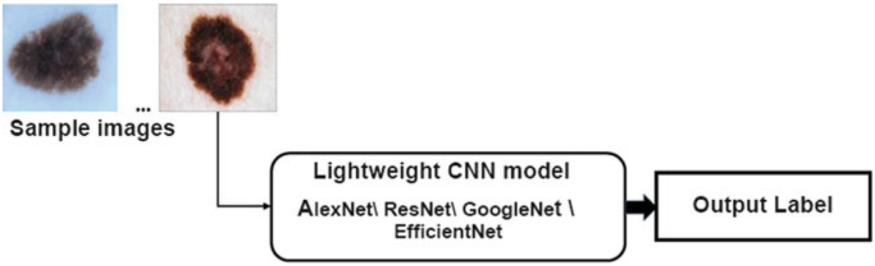


**Fig. 2** Proposed lightweight CNN model for skin cancer classification from dermoscopic images

**EfficientNetb0** EfficientNet is a convolutional neural network which performs scaling on depth, width, and resolution forming a series of networks that vary from b0 to b9. It is composed of 237 layers and can train more than one million images with less parameters. The baseline model EfficientNetb0 [11] is utilized in this work.

**ResNet50** ResNet50 is a convolutional neural network which is also known as residual network [12]. It uses skip connection concept. The residual blocks in the network can solve the problem of exploding gradients, and enhance the performance of neural network with more layers.

## 4 Experimental Study

The above mentioned CNN architectures are utilized to classify the sample images from ISBI image set [7]. It contains 1279 samples.

The pretrained deep network models GoogleNet, AlexNet, EfficientNetbo, and ResNet50 are utilized to classify sample images in the test set. The classification by GoogleNet model gives an accuracy of 81.11% . Similarly, the classification results using AlexNet, EfficientNetb0, and ResNet50 models give accuracies of 77.04%, 80.83%, and 76.11%, respectively. Table 1, shows the performance comparison of different CNN networks utilized in the proposed study for skin cancer classification. The analysis of accuracies of different networks shows that GoogleNet is giving the highest accuracy compared to other networks. Besides this, the GoogleNet cancer prediction is compared with previous works in Table 2, and it shows that GoogleNet has attained better classification performance than others.

**Table 1** Comparison of the performance of skin cancer classification using different lightweight CNN models

| LightWeight CNN | Accuracy (%) |
|---|---|
| GoogleNet | 81.11 |
| EfficientNet | 80.83 |
| AlexNet | 77.04 |
| ResNet | 76.11 |

**Table 2** Comparison of melanoma detection by GoogleNet model with the existing methods

| Author | Method | Accuracy (%) |
|---|---|---|
| Kawahara and Hamarneh [4] | CNN ensemble | 79.5 |
| Nasr-Esfahani [5] | 2 layer CNN | 81 |
| Proposed | GoogleNet | 81.11 |

## 5 Conclusion

Automatic detection of skin cancer will be very useful for patients as well as doctors. This paper presents a deep model to automatically detect skin cancer. Pretrained CNN models GoogleNet, AlexNet, EfficientNet, and ResNet have been used for building the cancer prediction models. The comparison of different networks for skin cancer classification has shown that GoogleNet has got better accuracy of 81.11% for the ISBI 2016 challenge dataset.

**Acknowledgements** Writers are grateful to the parent institution in getting help.

## References

1. Skin cancer foundation home page. https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/. Last accessed 2022/2/11
2. Jerant AF, Johnson JT, Sheridan CD, Caffrey TJ (2000) Early detection and treatment of skin cancer. Am Family Phys 62(2):357–386
3. Yuan Y, Chao M, Lo Y-C (2017) Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans Med Imaging 36(9):1876–1886
4. Kawahara J, Hamarneh G (2016) Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. In: Proceedings of the 7th international conference on machine learning in medical imaging (MLMI 2016)
5. Nasr-Esfahani E, Samavi S, Karimi N, Soroushmehr SMR, Jafari MH, Ward K, Najarian K. Melanoma detection by analysis of clinical images using convolutional neural network. In: Proceedings of the 38th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 16–20
6. Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, Alsaiari SA, Saeed AHM, Alraddadi MO, Mahnashi MH (2021) Skin cancer detection: a review using deep learning techniques. Int J Environ Res Public Health 18(10):5479
7. 2016 challenge dataset. https://challenge.isic-archive.com/data/. Last accessed 2022/2/8
8. Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53(8):5455–5516. https://doi.org/10.1007/s10462-020-09825-6
9. Szegedy C et al (2015) Going deeper with convolutions. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 1–9. https://doi.org/10.1109/CVPR.2015.7298594
10. Alom MZ et al (2018) The history began from AlexNet: a comprehensive survey on deep learning
11. Tan M et al (2019) EfficientNet: rethinking model scaling for convolutional neural networks
12. Ji Q et al (2019) Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images

# Collective Behavior in Community-Structured Network and Epidemic Dynamics

**Kirti Jain, Vasudha Bhatnagar, and Sharanjit Kaur**

**Abstract** Contagious diseases spread in population thru the contact network and their spread is a function of the complex interplay of the biological infectivity and behavior of individuals. In this research, we aim to understand how the epidemic dynamics is impacted by the collective behavior of individuals in communities. The stochastic block model is used to generate a community-structured network to investigate the spread of disease using the classical SIR spreading model. We model individual behavior as *fear-index* that indicates the extent to which an individual follows health and hygiene protocols as a self-protective measure against disease. We observe that the collective behavior of individuals in a community strongly influences the course of an epidemic. Infected individuals with low fear-index rapidly spread the infection within and outside the community. Furthermore, low fear in communities leads to faster growth of the epidemic. We also find that the communities, which comply with the restrictions manifesting high fear level, also suffer the burden of the disease because of non-compliance by other communities (low fear level). Communities with low fear levels are 'high risk' groups and should be targeted for awareness campaigns.

**Keywords** Epidemic dynamics · Contact networks · Community-structured networks · Fear-index · Collective behavior · Social norms · Seed infection

---

K. Jain (✉) · V. Bhatnagar
Department of Computer Science, University of Delhi, Delhi, India
e-mail: kjain1@cs.du.ac.in

V. Bhatnagar
e-mail: vbhatnagar@cs.du.ac.in

S. Kaur
Acharya Narendra Dev College, University of Delhi, Delhi, India
e-mail: sharanjitkaur@andc.du.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_16

## 1 Introduction

Computer virus spreading, rumors circulation, infectious disease spread, etc., are common examples of the diffusion process, where a contagion spreads in the population through contact between entities. *Contact networks* mimic the linkages among the entities and are useful for studying diffusion processes for a given spreading model. These studies are important since the diffusion of virus/rumors/infectious disease, etc., significantly impacts the physical, social, and mental health of individuals in the society [1]. Understanding the dynamics of the spread of contagion is important to prevent, control, and devise strategies for mitigation of the catastrophe that arises due to spread [2–4].

The structure and topological properties of contact networks have been explored extensively to observe and understand the spread of contagious diseases [5–7]. Community structure, which is the hallmark of the real-world contact network, plays a notable role in the spread of contagion in networks. A community symbolizes a group of individuals that display homophily; i.e., they bear similarities of interests, goals, or other attributes. Connections (linkages) between individuals within a community (intra-community connections) are more dense than the connections between individuals belonging to different communities [8].

Due to mesoscale description of communities in networks, their role in inhibiting and diffusing epidemic spreading has been explored earlier [4, 8–11]. However, studies investigating the extent to which the community structure affects epidemics is still under-explored [1, 8]. A mesoscopic scale description of the epidemic helps in implementing appropriate non-pharmaceutical interventions (NPIs) to constrain movement in the neighborhood to control the contagious infectivity spread with no known treatment like Nipah virus and COVID [8, 12]. Recently, Gosak et al. have shown that the effectiveness of imposing lockdown thru community structure increases if implemented jointly with social distancing among individuals across the whole network [3].

Behavior plays a crucial role in the spread of contagious diseases. It is also observed that people often change their behavior during an infectious disease outbreak under the influence of the society/group they belong to. Thereby, their risk of infection is also dependent on the behavior of their contacts along with their self-awareness [4, 13, 14]. The level of the perceived risk of infection varies greatly among different groups in the population that may be attributed to varying levels of information regarding the pathogen, quality of local healthcare, availability of preventative measures, etc. [9]. Societies undertaking precautionary advisories reduce susceptibility to the disease and thus potentially control the size of an epidemic outbreak by imposing movement restrictions outside the group. For this reason, ignoring the collective change in human behavior may adversely impact the accuracy of disease dynamics models. In previous meta-population epidemic dynamics studies, a single infection rate is used for the entire network. In contrast, in reality, human disease infectivity varies because of social, geographical, or cultural aspects within different subpopulations [14, 15].

### 1.1 Motivation and Contribution

Gosak et al. recently showed that individual awareness and social distancing reduce the count of infected individuals along with flattening of the infection curve compared to pure top-down approaches like community lockdown [3]. Motivated by this, we aim to understand the collective impact of primary emotion *fear* (that is expressed at three levels, viz. psychological, physiological, and behavioral [16]), prevailing in a community on epidemic dynamics and its role in preventing the infection spread. *Fear* is a conscious emotion that follows as a defensive behavior during the spread of disease for which no cure or vaccination is yet formalized [17]. The main objective is to show the importance of collective behavior in terms of prevailing fear emotion within a society that may trigger or control the disease spread and hence must be studied critically while projecting the epidemic variables.

We study the epidemic dynamics on synthetic homogeneous community-structured networks generated using the stochastic block model (SBM). A *fear-index* is used to model an individuals' behavior during disease spread and its impact on the risk of infection. We analyze epidemic dynamics in community-structured networks by modeling heterogeneous behavior in communities. We believe that collective behavior within society in terms of following protocols, maintaining social distancing, avoiding unnecessary meetings, awareness about disease prevalence reduces infectivity in the early stage of disease spread and may prevent an epidemic outbreak. Introducing fear dynamics at a community level and incorporating change in individuals' behavior during disease outbreaks results in realistic epidemic variables observed in the real world.

### 1.2 Organization

The structure of the paper is as follows. Section 2 briefly details the related work on epidemic spread in community-structured networks. Section 3 describes the methodology adopted in capturing collective behavior in a homogeneous community-structured network and details the impact of fear dynamics in communities on the epidemic simulated using a traditional spreading model. Experimental design is given in Sect. 4, and our findings along with the discussion on potential applications of the proposed model are presented in Sect. 5. Conclusion and suggestions for future work are given in Sect. 6.

## 2 Related Works

The presence of community structure in the network has a considerable impact on the epidemic dynamics [9, 18–20]. In a study investigating the spread of disease in a network with community structure, Kitchovitch et al. demonstrate that disease spreads more rapidly within the community than across communities [9]. Their work assumed that all individuals in a community bear the same and constant awareness level throughout the epidemic outbreak. Huang and Li studied the susceptible–infected (SI) dynamics on the scale-free network with communities. They found that the community structure inhibits the infection spread and reduces the epidemic prevalence [18]. Salathé and Jones investigated the epidemic spread in empirical and artificial networks and got a similar conclusion. They found that immunizing bridge nodes is more effective than simply immunizing highly connected nodes in community-structured networks [20]. Stegehuis et al. studied two random network models to create networks with communities and show that the exact internal structures of communities barely influence the behavior of the spreading process [8]. Liu et al. studied epidemic dynamics on two-layer networks with the assumption that individuals' awareness depends on the size of the community they belong to [19]. Silk et al. modeled adherence and concern of the disease at an individual level where adherent individuals cut their connections in the network [21].

In existing community-based epidemic modeling, the researchers ignored the behavioral response and changes in an individual's fear of disease that depends on the cultural preference of the community. The need of the hour is to incorporate changes in individuals' fear due to prevailing infections and recoveries as the epidemic progresses. Including such behavioral changes in epidemic modeling will result in realistic epidemic dynamics.

## 3 Methods

This section presents a detailed description of the model used to generate a network with the desired community structure, followed by a narration of the epidemic-spreading process used in this work. Subsequently, we explain the notion of the collective behavior in a community and its repercussions on the transmission probability of the disease and epidemic parameters.

### *3.1 Generation of Network with Community Structure*

Stochastic block model (SBM) has become increasingly popular in statistical analysis of graphs with clustering structure and high modularity [22]. SBM is a generative model for random graphs, which places $N$ vertices into $m$ communities such that pairs of nodes are connected with specified edge probabilities. There are dense connections between the nodes of a community and sparse connections between nodes in different communities.

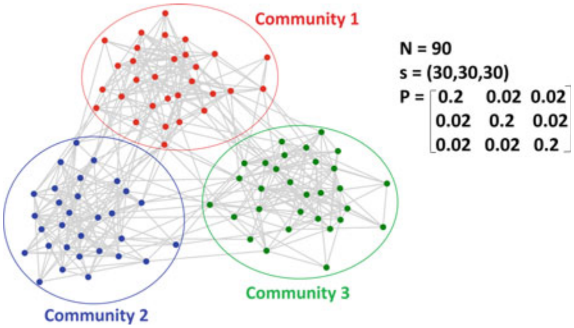


**Fig. 1** A toy homogeneous community-structured network with three equi-sized communities generated using stochastic block model for given community size vector $s$ and probability matrix $P$

**Fig. 2** Standard SIR model for disease spread

Let $n_i$ be the number of nodes (order) in community $C_i$ $(1 \leq i \leq m)$, and $N$ $(= \sum_{i=1}^{m} n_i)$ be the total number of nodes in the network. The SBM is defined by the triplet $(N, s, P)$, where $s = (n_1, \ldots, n_m)$ is a vector of size $m$, corresponding to $m$ communities in the network, and $P$ is a $m \times m$ symmetric matrix, where $p_{ij}$ is connection probability between nodes of communities $i$ and $j$. It is noteworthy that the diagonal elements of $P$ indicate the cohesion within the community, and off-diagonal elements indicate coupling density between two communities. If all probabilities in $P$ are equal, the community structure degenerates, and the network reduces to Erdos-Renyi random graph without any community structure. If all communities have same cohesion $(p_{ii} = p_{jj}\ \forall i, j)$ and the coupling density is same for all pairs of communities, then the model is termed as symmetric SBM [23]. We refer to equi-sized communities generated using symmetric SBM as *homogeneous community structure*. Figure 1 shows a toy network of order $N = 90$, with three homogeneous communities generated using symmetric SBM with given community size vector $s$ and probability matrix $P$.

## *3.2 Spreading Model*

In this work, we adopt the classical susceptible–infected–recovered (SIR) model for understanding the influence of collective behavior of communities on the dynamical process of disease spreading (Fig. 2). At any time, an individual in the network may be in one of the following states, viz. susceptible, infected, and recovered. Any susceptible individual in contact with an infected individual gets infected with probability $\beta$ (transmission probability). Each infected individual recovers with probability $\gamma$ and remains permanently immunized to the infection.

### *3.3 Collective Behavior in a Community*

By definition, individuals in a community are more similar to each other, and their behavior is influenced by that of other members. We model individual behavior in terms of *fear-index* that represents an individual's fear for the disease. The fear-index encompasses the impact of multiple extraneous factors like awareness of a disease, public health education, personal relevance and risk, information dissemination by news media, religious and cultural congregations [13]. We list below the heuristics employed for modeling fear in the communities.

**Individual fear**: During simulation, we associate a *fear-index* with each individual in the population to quantify her illness anxiety disorder and to characterize the behavioral response to the prevalent disease. Individuals follow the advice and stringent rules set by governments and bear all precautionary responses in accordance with their fear level.

At time $t$, an individual $u$ has fear-index $\phi_t^u$ $(0 \leq \phi_t^u \leq 1)$. The value of $\phi_t^u$ closer to 1 corresponds to a person's high fear of the disease, which causes them to behave extremely cautiously (follow hygiene protocols, social distancing, etc.) to reduce the risk of getting infected. Conversely, a fear-index closer to 0 indicates the casual attitude and reluctance to follow norms, thereby increasing the probability of infection. The fear-index of an individual is dynamic and varies with time in accordance with infectivity and recovery in the neighborhood. Thus, at any time, an individual has two variables, viz. health state and fear-index. Figure 3 shows a snapshot of a toy network of seven individuals in different states of the disease using the SIR model along with their respective fear-index values. Susceptible individuals ($U$, $X$, and $Z$) have medium fear, while $V$ and $W$ have contracted the disease and have higher fear. Individuals $T$ and $Y$ possess the lowest fear due to recovery from the disease.

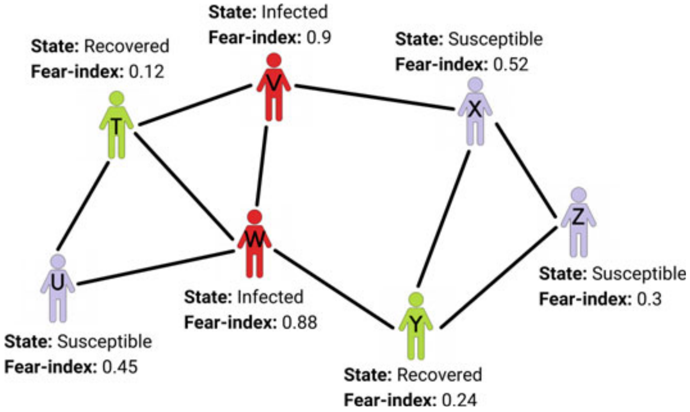


**Fig. 3** Toy network of seven individuals showing their SIR states and fear-index values

**Initial fear in community**: Disparities in societal cultures and innate psychological differences vary with communities and manifest the fear-index of an individual. The collective behavior of individuals in a community at the onset of disease influences the epidemic spread. Initially, collective fear in community $C_i$ is normally distributed ($N(\mu, \sigma)$) among individuals with mean $\mu$ and standard deviation $\sigma$. Individual heterogeneity in fear is controlled by the value of $\sigma$ regardless of her state.

We categorize the initial average fear $\mu_0^i$ in community $C_i$ at time $t_0$ as {*very-low*, *low*, *medium*, *high* and *very-high*} using the distribution $\{N(0.1, 0.025)$, $N(0.3, 0.025)$, $N(0.5, 0.025)$, $N(0.7, 0.025)$, and $N(0.9, 0.025)\}$, respectively. Our hypothesis is that the dispersion of epidemic spread is dependent both on the community structure and mean behavior of individuals in the community. Highly aware subpopulation suppresses the spread over the lackadaisical group by slowing down the transmission probability of the contagion.

**Fear dynamics among individuals and their neighbors**: Although we set the initial collective behavior of a community, the fear of underlying individual changes with time during disease spread. Fluctuating fear is attributed to the health condition of an individual and her neighbors, awareness about the disease, quality of medical infrastructure, rate of spread of infection, participation in mass gatherings, etc. Individuals suffering from the disease usually experience high fear and transmit it to their social contacts. When an individual recovers, her fear decreases as she feels relieved. When an individual $u$ transits to $I$ or $R$ states, fear-index $\phi_t^u$ is reset to *very-high* ($N(0.9, 0.025)$) and *very-low* ($N(0.1, 0.025)$), respectively. We assume that the recovered individuals possess some fear and anxiety, though reduced.

In view of fear as a transmissible contagion, we update the fear-index ($\phi_{t+1}^v$) of a susceptible neighbor $v$ of infected or recovered individual $u$ by taking *average* of their fears. Updating fear using average is interpretable and reflects the real-world situation. Fear rises when direct contacts of individuals suffer illness and declines when the contacts recover.

$$\phi_{t+1}^v = \frac{\phi_t^v + \phi_t^u}{2}, \quad \forall v \in \mathrm{Nbr}(u) \tag{1}$$

where $\mathrm{Nbr}(u)$ is the set of susceptible neighbors of individual $u$.

**Fear-based transmission probability**: Conventional models for epidemiological studies consider a uniform probability of infection ($\beta$) for all individuals in the population. This assumption is at odds with the real-world scenario, where individual behavior influences the probability of contracting the disease. Highly cautious individuals protect themselves from the infection and slash down the chances of infection. At the same time, a casual attitude toward hygiene and safety protocols enhances the probability of contracting the disease. Thus, the probability of contracting a disease is governed by individual behavior and is variable in the real world. As the epidemic progresses, individual behavior (fear) may change due to the state and neighbors' behavior. Thus, the probability of infection is a characteristic of an individual and varies within a community.

The number of infected neighbors increases the likelihood of an individual getting infected. Therefore, the perception of risk of a susceptible individual $u$ at time $t$ depends on the number of infected neighbors $(i)$ and the fear of disease $(\phi_t^u)$. The effective probability $(\hat{\beta_t^u})$ of a susceptible individual $u$ getting infected is a function of individual behavior and is computed as:

$$\hat{\beta_t^u} = \beta^{(\phi_t^u)^2} * e^{-\frac{1}{i}} \tag{2}$$

where $\beta$ is the base transmission probability of the disease, $i$ is the number of infected neighbors, and $\phi_t^u$ is the fear of susceptible individual $u$ at time $t$. Incorporating the fear level of an individual to compute infectivity before the state transition from susceptible to infected will not only capture the individual's behavior but will also aid in projecting realistic epidemic variables.

## 4 Experimental Design

Earlier studies on random graph models with community structure have shown that community structure influences the epidemic-spreading process [9, 10, 20]. We extend the study to community-structured networks and investigate the effect of collective behavior of a community on the epidemic dynamics. We design experiments to answer the following questions.

1. *How does the collective behavior of communities in a population influence the spreading process?*
   Individuals in communities with high fear levels tend to follow health-related protocols to slow down the infection spread. On the contrary, the casual attitude of individuals in a community with low fear levels enhances the probability of contracting the disease. We show the progression of infected and recovered cases in the community-structured network with varying levels of fear. We also study the role of community structure and fear of disease in population on the epidemic dynamics by comparing the number of infected cases in networks with and without community structure (Sect. 5.1).
2. *How does the interplay of transmission probability of the disease and fear in the population impact the number of infections?*
   Increased fear in the population retards the rate of epidemic spread, while an increase in transmission probability fuels the growth. We vary the base transmission probability and fear in a community and observe the effect of collective behavior on peak infections (Sect. 5.2).

3. *How does the fear level of the initially infected node (individual) influence the infection spread?*
   We vary fear of the initially infected node to show the alteration caused by it on the infection spreading. We also compute the average time for the infection to reach outside the community with varying levels of fear to show its influence on disease spread (Sect. 5.3).
4. *What is the impact of different levels of fear in communities on the infection spreading ?*
   Individuals in communities with high fear levels tend to follow health-related protocols to reduce disease transmission. We model a population with communities having different levels of fear and observe community-wise infection spread (Sect. 5.4).

We generate a homogeneous community-structured network of 2000 nodes using symmetric stochastic block model with the following parameters.

$$\begin{aligned} s &= (400, 400, 400, 400, 400) \\ P &= \begin{bmatrix} 0.02 & 0.001 & 0.001 & 0.001 & 0.001 \\ 0.001 & 0.02 & 0.001 & 0.001 & 0.001 \\ 0.001 & 0.001 & 0.02 & 0.001 & 0.001 \\ 0.001 & 0.001 & 0.001 & 0.02 & 0.001 \\ 0.001 & 0.001 & 0.001 & 0.001 & 0.02 \end{bmatrix} \end{aligned} \tag{3}$$

The resulting network has 9629 edges ($M$), average degree ($k$) as 9.62, and modularity as 0.592. In the rest of the paper, unless otherwise specified, we use base transmission probability $\beta = 0.1$ and recovery rate $\gamma = 0.2$ for all the simulations adopting the SIR model [20]. In our study, we use relatively small value for $\beta$, for which $R_0 \approx (\frac{\beta}{\gamma}) * k \approx 4.8 > 1$, a known condition for epidemics.

We select the most central node in a community using eigenvector centrality [5]. This node is set as the initially infected node, and the remaining individuals of the network are susceptible. We refer to the community from where the infection starts as the *seeded community* and the first infected node as *seed node*. In all the experiments except the last two (Sect. 3.3), initial fear in the seeded community is varied, and fear in other communities is set to *medium*. All the reported results are averaged over 100 simulations.

## 5 Results and Discussion

This section details the results of experiments conducted to answer the questions raised in the previous section.

## *5.1 Collective Behavior in Community and Epidemic Dynamics*

This experiment aims to study the influence of collective behavior in a community on epidemic dynamics. We visualize the spread of epidemic in community-structured network for two extreme initial fear levels (*very-low* and *very-high*) of the seeded community ($C_1$). *Low* initial fear showcases the casual attitude of individuals toward the health protocols, while *very-high* initial fear induces controlled and self-protective behavior within the community. Other communities ($C_2 - C_5$) have same initial fear level (*medium*). The simulation begins at time $t = 0$, with the most central node in $C_1$ as the seed node. Figure 4a, b shows the snapshots of epidemic spread in *five*-community network with *very-low* and *very-high* initial fear levels, respectively, in $C_1$.

*Very-low* fear of individuals in community $C_1$ (violet color) results in a faster course of epidemic in panel (a) compared to panel (b) where the seeded community has *very-high* fear level. We also observe that the infection is largely confined to the seeded community during the early stage of spread ($t = 5$) for both cases. A high number of infections in panel (a) is attributed to *very-low* fear of individuals in the

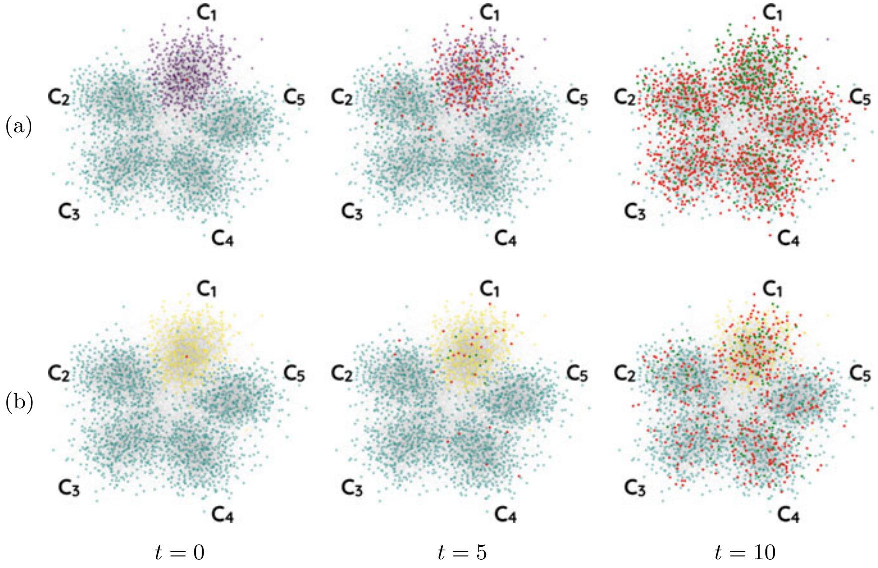


**Fig. 4** Panels **a** and **b** show the snapshots of a community-structured network during the course of epidemic for two initial fear levels, viz. *very-low* and *very-high* in community $C_1$ (nodes marked in violet and yellow colors, respectively, at $t = 0$). Other communities initially have *medium* fear (nodes marked in bluish-green color). Susceptible individuals are marked in colors of their respective communities, whereas infected and recovered individuals are marked with red and green colors, respectively

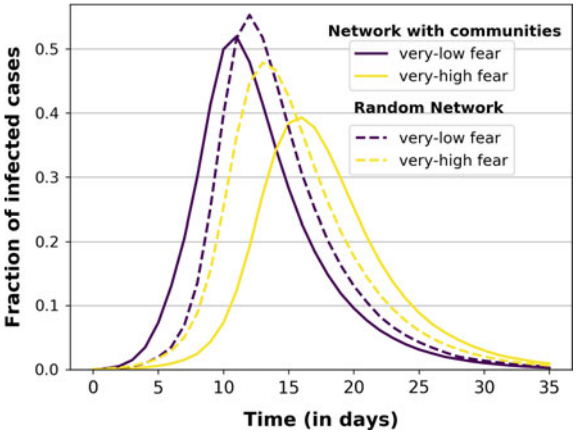


**Fig. 5** Infected cases in community-structured network (solid line) and random network (dash line) for two fear levels viz. *very-low* and *very-high*

community, which elevates the probability of infection. As the epidemic progresses ($t = 10$), the infection starts spreading to other communities, and the number of infected individuals (red-colored nodes) is comparatively higher in panel (a). The slower spread of infection in the network is attributed to initially *very-high* fear in panel (b).

Next, we aim to study the difference in epidemic dynamics in networks with and without community structure. We construct a random network of the same size and order as the community-structured network ($N = 2000$, $M \approx 9629$). We randomly pick a sample of 400 nodes (size of seeded community) from the network and set their initial fear level as *very-low* and *very-high* in two different simulations. The remaining 1600 nodes are assigned *medium* fear level, and the spread is initiated from the central seed node in the network. We plot the fraction of infected cases for the community-structured and random networks in Fig. 5.

It is indicated from the figure that *very-low* collective fear of individuals in the seeded community results in rapid infection spread (solid violet line). Since individuals in a community are densely connected, the infection spreads quickly within the community before propagating to other parts of the network. Low collective fear in a group flares up the individual transmission probability resulting in a larger outbreak. When the infection spreads to other communities (with *medium* average fear), it is suppressed due to comparatively more cautious behavior of individuals in these communities. In contrast, in random network, individuals with *very-low* fear are distributed randomly in the population, and most of their neighboring individuals have *medium* fear of disease. As a result, the infection spread is slightly delayed, though the peak is comparatively higher (dash violet line). Thus, initial *very-low* fear in the seeded community leads to faster growth of the contagion followed by its mitigation in a community-structured network compared to the random network.

In the second scenario, *very-high* fear level of individuals in $C_1$ makes it harder for the infection to spread from the localized group, reducing the overall infection curve (solid yellow line). This is also observed from Fig. 4. When infection propagates to other communities with comparatively low fear, the growth rate increases, but these infections can no longer cause a notable outbreak. Contrastingly, in random network, individuals with high fear levels scattered in the population are not effective in containing the epidemic despite their self-protective individual behavior. As a result, more cases are reported in random network than in community-structured network. This illustrates the importance of collective behavior in understanding epidemic dynamics.

*We believe that the collective behavior of like-minded individuals in a community considerably impacts the spreading process and must be considered while projecting realistic epidemic variables.*

## 5.2 Initial Fear and Transmissibility of the Pathogen

Epidemic dynamics change drastically due to the complex intertwining of the societal fear and transmissibility of the pathogen. As individual fear emotion varies with time, the effective disease transmission probability for an individual also changes (Eq. 2). In this experiment, we show the sensitivity of the transmission probability of the pathogen to the initial collective fear in the community. We vary the base transmission probability in the range 0 to 1 in steps of 0.01, which models the differences in transmissibility of different contagious diseases. We also vary the initial fear level in the seeded community ($C_1$) in steps of 0.1, note the peak number of cases, and plot the heatmap in Fig. 6.

It is clearly visible from the heatmap that the peak cases decrease with increasing initial fear in the seeded community for all values of $\beta$. For a fixed fear level in the community, the number of peak cases increases with the pathogen's transmissibility ($\beta$). The top-left region of the heatmap shows the maximum cases are reported

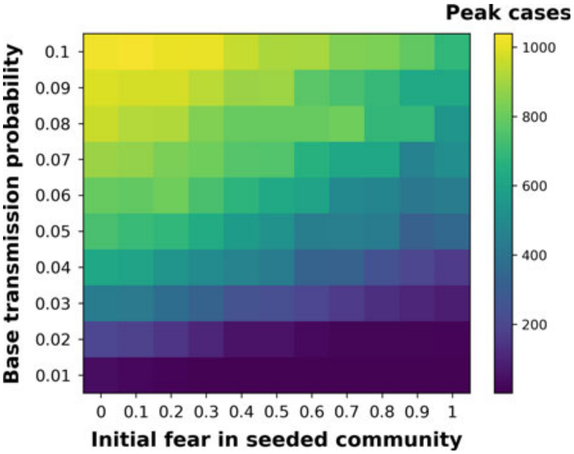


**Fig. 6** Heatmap of total peak cases with varying initial fear (in seeded community) for increasing values of base transmission probability ($\beta$) in the homogeneous community-structured network

when the infectivity of the disease is high and fear in the seeded community is low. The bottom-right region shows that when the pathogen is feeble and fear among individuals is high, peak cases are minimum.

*Results of this experiment show that the fear of disease and transmissibility of pathogen act as two opposing forces on the epidemic spread. Low fear in the community and high transmission probability fuels the infection spread and leads to an uncontrolled outbreak. In case the pathogen is highly transmissible, even high collective behavior cannot confine the rapid spread of infection in the population.*

## 5.3 Role of Fear of Seed Node on Infection Propagation

We conjecture that fear of the seed node plays an important role in controlling the spread. To validate this conjecture, we vary the fear of seed node in steps of 0.1 while setting the initial fear of all the communities to *medium*. As the average path length of each community is approximately 3, the infection is expected to remain mostly within the seeded community till $t = 3$ and start propagating outside the community at $t = 5$. Figure 7 shows the fraction of total infected cases in the network at two time instances $t = 3$ (orange bars) and $t = 5$ (blue bars). The lines of the same color indicate the infected cases in the seeded community with varying fear of infected seed node.

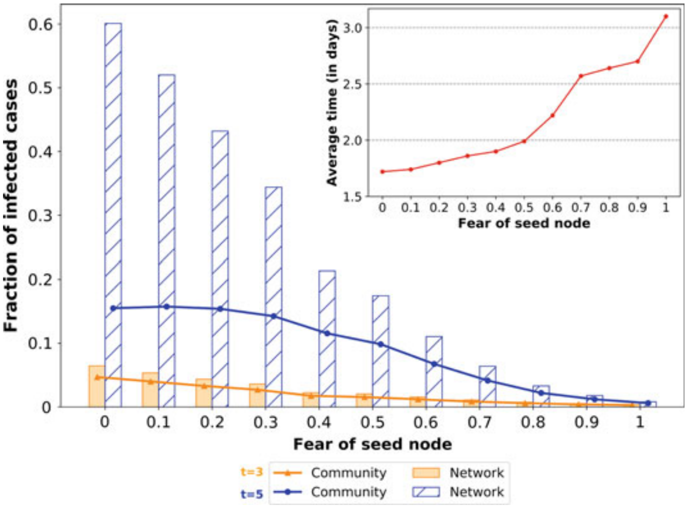


**Fig. 7** Fraction of total infected cases at time $t = 3$ (orange bars) and $t = 5$ (blue bars), with varying fear of the seed node. Lines of same color indicate the fraction of infections in the seeded community. Inset figure shows average time (in days) for infection to reach outside the seeded community for varying fear of the seed node

The figure shows that low fear of the seed node (individual) results in a higher fraction of cases in the early days of the epidemic (orange bars). With an increase in *fear* of the seed node, the transmission probability decreases, due to which the fraction of infected cases reduces.

Since the average path length of the seeded community is $\approx 3$, most of the infection spread is contained within the community (orange line) at $t = 3$ with negligible infected cases outside the seeded community (orange bar). However, at time ($t = 5$), the infection spreads to other communities leading to an overall higher fraction of infected cases (blue bar). The spread to other parts of the network reduces sharply with an increase in fear of the seed node (blue line).

It is important to detect the outbreak as early as possible in applying mitigation strategies. There is a noticeable period before an outbreak in the seeded community reaches the rest of the population. Inset in Fig. 7 shows the average time for the infection to reach outside the seeded community. It is observed that when seed fear is low, the infection reaches quickly ($\approx$ two days) to other communities. This is attributed to an increase in the effective transmission probability due to low fear. On the other hand, an increase in fear of seed node slows down the spread within the seeded community, and hence time to reach the first infection outside that community increases.

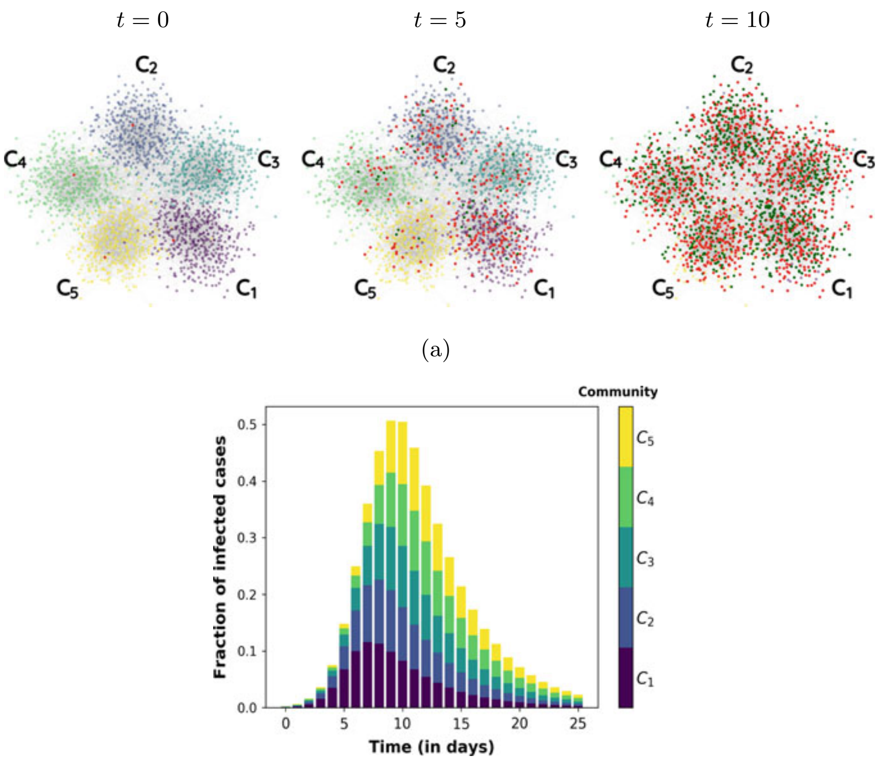
*Thus, fear of an infected individual directly impacts the spreading rate and determines the course of disease spread. Hence, it is crucial for containing the infection in the confined area and is an important determinant of the overall prevalence of the disease.*

### 5.4 Impact of Varying Collective Behavior on Community-Wise Infection Spread

Non-pharmaceutical interventions imposed by local administrations have a marked impact on the collective behavior. Using fear level of the community as a proxy for collective behavior, we assess the effectiveness of the enforced restrictions in controlling the epidemic dynamics. Figure 8a shows the progression of epidemic in the network.

We model different types of community behavior by setting different initial fear levels for the five communities in the network. We set the initial fear in communities $C_1 - C_5$ as {*very-low*, *low*, *medium*, *high* and *very-high*}. The highest fear level in a community portrays the strictest following of the imposed restrictions, while the lowest fear level models the most casual collective behavior of the individuals in the community. We select the seed node in each community based on the centrality and simulate the SIR spreading model. We note community-wise infections each day.

*Very-low* fear of individuals in community $C_1$ (violet color) results in the fastest course of epidemic, while community $C_5$ (yellow color) with *very-high* fear level has least number of infections at $t = 5$. We observe that the infection is largely confined



(a)

(b)

**Fig. 8** **a** Snapshot of epidemic in community-structured network with initially infected node (red color) in each community. **b** Comparison of infections in communities $C_1 - C_5$ with different initial fear levels {*very-low*, *low*, *medium*, *high* and *very-high*}

within each community during the early stage of spread. The high number of infections in community $C_1$ is attributed to *very-low* fear of individuals, which elevates the probability of infection. Since communities with low fear induce infection to other communities, a nearly equal number of infections are observed in all communities at time $t = 10$. Figure 8b shows the stacked bar of community-wise infections reported on each day. The results are in accordance with the dynamics observed in Fig. 8a.

*The results illustrate that the communities which comply with the restrictions (i.e., high fear level) also suffer the burden of the disease because of non-compliance by other communities (low fear level). Communities with low fear levels are 'high risk' groups and should be targeted for awareness campaigns. These groups should also be prioritized for vaccination and other preventative measures for the well-being of the population at large.*

## 6 Conclusion and Future Work

Introducing collective behavior in terms of fear emotion within homogeneous community-structured networks show that epidemic dynamics changes with the varying fear level in communities. High fear of disease mitigates the disease spread, whereas low fear fuels the propagation by impacting the underlying transmission probability of the contagion. We demonstrate that fear of the infected seed node can significantly alter epidemic dynamics. Also, the time to reach infection outside the seeded community delays with increasing fear of the seed node and is a critical determinant for controlling the disease outbreak within the population. We intend to apply the proposed fear model on a heterogeneous community-structured network to improve its applicability and effectiveness.

## References

1. Wang S, Gong M, Liu W, Wu Y (2020) Preventing epidemic spreading in networks by community detection and memetic algorithm. Appl Soft Comput 89:106118
2. Chen Y, Paul G, Havlin S, Liljeros F, Stanley HE (2008) Finding a better immunization strategy. Phys Rev Lett 101:058701
3. Gosak M, Duh M, Markovič R, Perc M (2021) Community lockdowns in social networks hardly mitigate epidemic spreading. New J Phys 23(4):043039
4. Xu Z, Li K, Sun M, Fu X (2019) Interaction between epidemic spread and collective behavior in scale-free networks with community structure. J Theoret Biol 462:122–133
5. Barabási AL, Pósfai M (2016) Network science. Cambridge University Press, Cambridge
6. Keeling MJ, Eames KTD (2005) Networks and epidemic models. J Roy Soc Interface 2(4):295–307
7. Small M, Cavanagh D (2020) Modelling strong control measures for epidemic propagation with networks—a COVID-19 case study. IEEE Access 8:109719–109731
8. Stegehuis C, Van Der Hofstad R, van Leeuwaarden JSH (2016) Epidemic spreading on complex networks with community structures. Sci Rep 6:29748
9. Kitchovitch S, Liò P (2011) Community structure in social networks: applications for epidemiological modelling. PLoS One 6:1–17
10. Liu Z, Hu B (2005) Epidemic spreading in community networks. Europhys Lett 72(2):315–321
11. Yu Y, Xiao G (2017) Infection spreading, detection and control in community networks. J Complex Netw 5(4):625–640
12. Prasse B, Devriendt K, Van Mieghem P (2021) Clustering for epidemics on networks: a geometric approach. Chaos Interdiscip J Nonlinear Sci 31(6):063115
13. Funk S, Salathé M, Jansen V (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. J Roy Soc Interface 7:1247–1256
14. Gelfand M, Jackson J, Pan X, Nau D, Pieper D, Denison E, Dagher M, Lange P, Chiu CY, Wang M (2021) The relationship between cultural tightness-looseness and covid-19 cases and deaths: a global analysis. Lancet Planet Health 5(3):e135–e144
15. Gong YW, Song YR, Jiang GP (2014) Epidemic spreading in metapopulation networks with heterogeneous infection rates. Phys A Stat Mech Appl 416:208–218
16. Steimer T (2002) The biology of fear-and anxiety-related behaviors. Dialogues Clin Neurosci 4(3):231
17. Coelho CM, Suttiwan P, Arato N, Zsido AN (2020) On the nature of fear and anxiety triggered by covid-19. Front Psychol 11:581314

18. Huang W, Li C (2007) Epidemic spreading in scale-free networks with community structure. J Stat Mech Theor Exp 2007(1):P01014
19. Liu T, Li P, Chen Y, Zhang J (2016) Community size effects on epidemic spreading in multiplex social networks. PLoS One 11(3):e0152021
20. Salathé M, Jones JH (2010) Dynamics and control of diseases in networks with community structure. PLoS Comput Biol 6(4):e1000736
21. Silk MJ, Carrignon S, Bentley RA, Fefferman NH (2021) Improving pandemic mitigation policies across communities through coupled dynamics of risk perception and infection. Proc Roy Soc B 288(1955):20210834
22. Holland PW, Laskey KB, Leinhardt S (1983) Stochastic blockmodels: first steps. Soc Netw 5(2):109–137
23. Abbe E (2017) Community detection and stochastic block models: recent developments. J Mach Learn Res 18(1):6446–6531

# Incorporating Semantics for Text Classification in Biomedical Domain

**Shikha Verma and Aditi Sharan**

**Abstract** Text classification is an important task with its scope in many applications like biomedical domain, bioinformatics, text mining, information retrieval, etc. Different methods have been proposed in literature for classification of textual data into binary classes. Multi-label text classification on the other hand becomes a real challenge when the number of classes is high in number or even when the classification aims for higher accuracy. The complexity grows when the content is morphologically rich and yet the solutions to them are eagerly sought after. For decades, there have been extensive discussions on how to select the relevant features in text analytics or text mining. The paper proposes 3 categories for the selection of relevant features in biomedical text, namely statistical features, biomedical features and linguistics features to categorize the abstracts of biomedical literature. The proposed model implements the SVM classification algorithm on the dataset of Hallmark of Cancer with more than 1800 abstracts taken into consideration.

**Keywords** Text classification · Multi-label classification · Biomedical dataset · Support vector machines · Statistical features · Biomedical features and linguistics features

## 1 Introduction

Natural language processing includes a common task of classification of text which assigns some predefined labels to the entire text. Since volumes of data especially the text data on the web are increasing rapidly, the only feasible solution for structuring the data and its metadata is by assigning labels automatically which eventually saves time and also the cost for these tasks.

---

S. Verma (✉) · A. Sharan
School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India
e-mail: shikha51_scs@jnu.ac.in

A. Sharan
e-mail: aditisharan@mail.jnu.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_17

The study of biomedical text data is increasing rapidly. With volumes of medical articles, the task of mining them and classifying these data appropriately is becoming a challenging task. A large number of biological information are being generated every year. Handling the biomedical information and then retrieving the relevant data from these sources are really important. Many a times even a single medical document like the patient discharge summary may consist of identification of multiple diseases or symptoms which needs to be classified accurately.

When it comes to information retrieval systems, PubMed [1, 2] is one of the most well-known biomedical databases, with over 20 million biomedical articles quoted from MEDLINE and Life Science publications as well as a user-friendly web-based search portal and an API for developers.

Multi-label classification is considered an effective method in assigning a suitable subset of labels to a given text document. Text classification in any natural language processing involves classifying the labels in different categories. These categories can be learned using the traditional supervised learning methods which involve the paradigm of classification of the text by converting the text to features. In binary classification, the sample or example under study with its feature vector is assigned only a single label. An important assumption here is that only one label and one semantic meaning is allotted to each training or test example. Binary classification task also considered the pattern of assigning labels to the examples within a subset of only two categories. In multiclass problem, it involves classification task where the labels are multiple, however, only one label is assigned as the output to the example as in the case of binary classification task. Multi-label classification on the other hand considered the assignment of multiple labels to each example which is practically the need for major classification tasks in many real life examples. For instance, a medical document like the patient abstract may consist of multiple disease or symptoms which can be further categorized into multiple diseases which will eventually lead to labeling with multiple classes at one go.

The goal of the division of multiple label classification is to predict a label set of test sample data provided from a given label set. Jin and Ghahramani [3] call these as semi-supervised problems, where each model is associated with more than one category, but only one of those classes is a true model category.

One of the new areas of research in the field of biomedicine is biomedical text mining. In fact, the growth of data/information in the field of biomedical literature is exponentially increasing day by day be it PubMed or the Medline Literature which has enormous volumes of data which is used by researchers but with such velocity of data in order to manage it should be mined and important information should be extracted from these biomedical knowledge.

The amount of biological literature is relatively vast and is rising rapidly, and the task of information retrieval is becoming more challenging. Text classification is one of the most challenging and typical tasks in supervised machine learning [4]. One of the most important and primitive issues in text classification is text preprocessing and selecting appropriate features that are appropriate representative of the text class. In this paper, we propose a semantic model of text classification using popularly using SVM classifier. Mostly, researchers performing text classification are using statistical features such as stemmed tf-idf vector of words or a combination of unigrams and bigrams. In this paper, we suggest the use of enhanced set features that can be helpful in incorporating semantics in classification specifically in context of biomedical domain. We have categorized the features into three categories: statistical, linguistics and biomedical features. We have done exhaustive experiment using SVM classifier to see the effect of hybridizing these features in various combinations. The results are quite motivating and suggest substantial improvements in traditional n-gram-based SVM.

The rest of the paper is organized as follows. Section 2 elaborates the role of SVM in text categorization and multi-label classification problem. Section 3 presents benchmark dataset and Sect. 4 elaborates the proposed model to classify biomedical abstracts. Section 5 explains feature extraction and categorization. The experimental design is highlighted in Sect. 6, and the study's findings are examined. Finally, Sect. 7 concludes the work by outlining future research directions.

## 2 Role of SVM and Multi-Label Learning

### *2.1 Support Vector Machine (SVM)*

The supervised machine learning algorithm: SVMs are gaining momentum in the machine learning community. This determines the boundary between the sample points in order to clearly distinguish the different classes. There have been many studies which focuses on the SVM algorithm works with the textual data. Other common automated text classification algorithms are $K$-Nearest Neighbor Method (KNN), Vector Space Method (VSM), Boosting Method, Artificial Neural Network Method, Naive Bayes and decision tree [5].

Among the other classification methods, SVM performs exceptionally well in case of the text data, and recently, there have also been researches focusing on the ranking of the SVM algorithm to work in the multi-label classification problems. Many practical applications, such as the handwritten text identification, social media monitoring, detecting trends in customer feedbacks, language detection, language translation and many more.

SVM has the following advantages:

- Solve machine learning problems with small sample cases
- Has excellent ability to generalize
- Helps break the curse of dimensionality
- Nonlinear problems can be easily solved with kernel-based methods.

## 3 Dataset

The experiment has been performed on the benchmark dataset named the Hallmark of Cancer dataset consisting of 1852 PubMed articles which are annotated text Baker et al. (2016). Since it is the multi-label classification problem, the abstract may be labeled with more than one label.

The ten hallmarks are given as below [6]:

- Sustaining proliferative signaling
- Evading growth suppressors
- Resisting cell death
- Enabling replicative immortality
- Inducing angiogenesis
- Activating invasion and metastasis
- Genome instability and mutation
- Tumor promoting inflammation
- Deregulating cellular energetics
- Avoiding immune destruction.

## 4 Proposed Model

The first step of the proposed model is preprocessing the 1852 PubMed articles which cleans the abstracts and removes the stop words, punctuations and other filtering then as a part of the feature extraction the proposed model has extracted the following features:

- Statistical features
- Linguistic features
- Biomedical features

All these features are explained in detail in the next section. After this stage, the SVM classifier is applied to above features and the model also hybridized the statistical features with linguistic and biomedical features. For experimental purpose, we have used 80% train and 20% test data. Section 5 discusses the experimental results in detail.

## 5 Feature Extraction and Representation

Because textual data is typically unstructured, feature extraction is an important component of machine learning systems. We need to build specific structural vectors from these datasets for machine learning algorithms to process. Feature extraction is a crucial stage in text categorization since the recovered features are utilized to represent texts. To extract relevant characteristics from textual information, a variety of approaches, feature models and algorithms are required.

In this paper, we consider a rich set of features that have been classified in three categories statistical, linguistic and biomedical features.

Accordingly, this section is divided in three subsections: statistical, linguistic and biomedical features, respectively (Fig. 1).

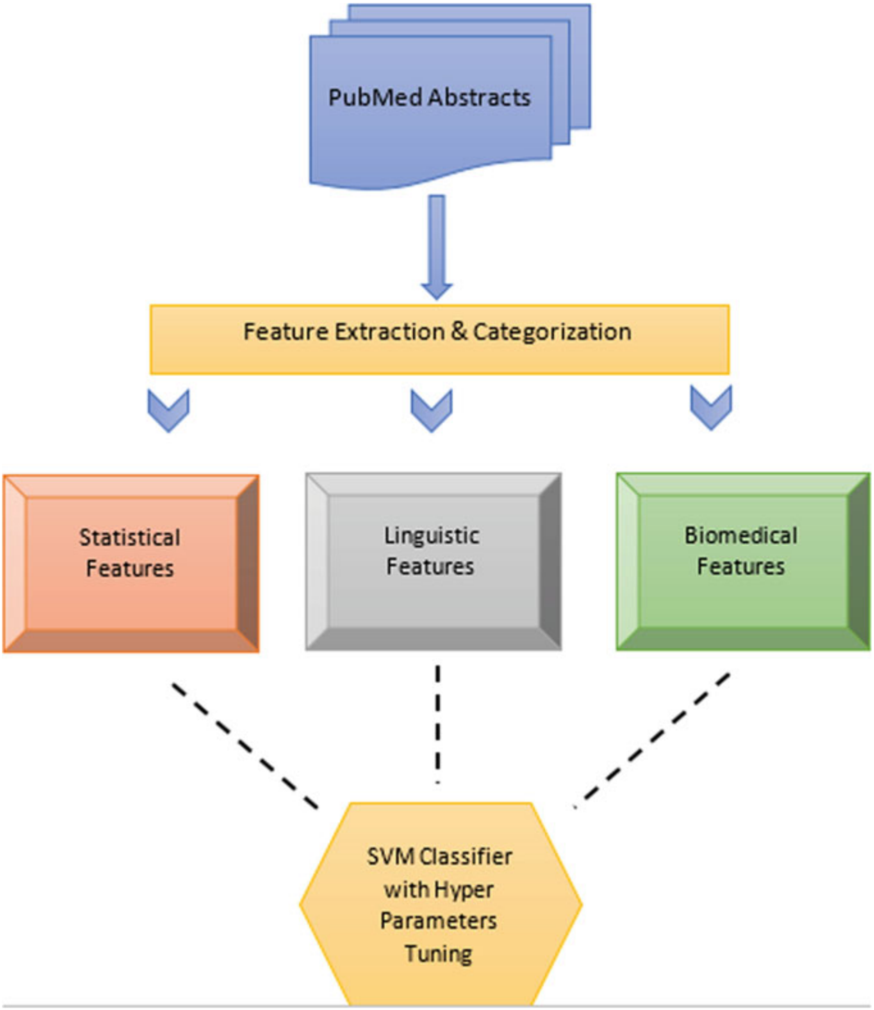


**Fig. 1** Layout of the proposed model

## *5.1 Statistical Features*

### N-gram features

The n-gram theory is used in this paper. A contiguous series of n tokens from a given sequence of a sentence is referred to as an n-gram. In most cases, an item is referred to as an atomic unit. A “unigram,” a “bigram” or a “trigram,” respectively, is an n-gram comprising one item, two items or three items. To extract the character n-gram from a document, you need to move the n character wide window one character at a time throughout the document. Traditional n-gram works on the word bigram model. We have used the noun bigrams in the proposed model to extract the noun bigrams as one of the feature to add richness. Here, when the abstract is tokenized into sentences, then the tokens of slice-2 are extracted which comprise the nouns. In this paper, we use the noun phrases as a bigram slice. Figure 2 shows the noun bigram slice being extracted as a feature from one of the abstract from Hallmark of Cancer dataset after tokenizing the paragraph into sentences and preprocessing the text.

### Lemmatized bag of words

A lemma [7] (derived from the Greek word “lemma,” “headword”) is the “dictionary” or “canonical” version of a group of words in linguistics. A lemma is the canonical form of a lexeme, where lexeme refers to the set of words with the same meaning and the word lemma refers to the representation of base word notation for that lexeme [6]. The objective of lemmatization is to find the lemma, or dictionary form, of a given word form. Here, we have used the lemmatized bag of words as a feature, and the Fig. 3 shows how the words in the sentence from the abstract are transformed into the lemma forms such as the word like “wanted” converge to want and the word “cells” converge to “cell.” So basically, this feature gives importance to the morphological analyzes of the word forms. Also note that the lemmatization has been applied after preprocessing of the input sentences.

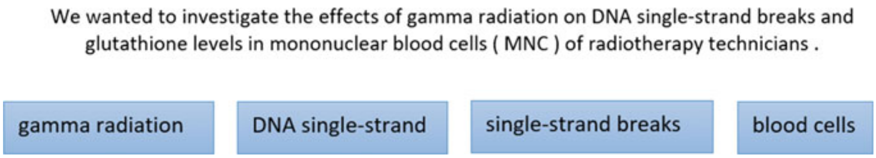
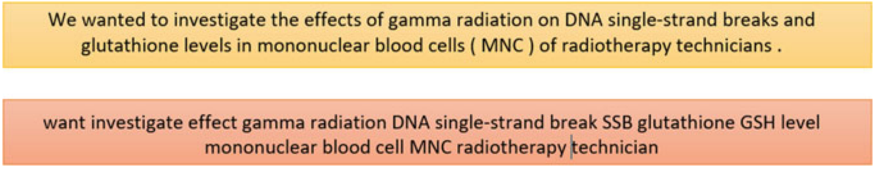

We wanted to investigate the effects of gamma radiation on DNA single-strand breaks and glutathione levels in mononuclear blood cells ( MNC ) of radiotherapy technicians .

**Fig. 2** Noun bigrams extracted from the abstract from hallmark of cancer dataset



We wanted to investigate the effects of gamma radiation on DNA single-strand breaks and glutathione levels in mononuclear blood cells ( MNC ) of radiotherapy technicians .

want investigate effect gamma radiation DNA single-strand break SSB glutathione GSH level mononuclear blood cell MNC radiotherapy technician

**Fig. 3** Lemmatized bag of words extracted from a sentence from hallmark of cancer dataset abstract

## *5.2 Biomedical Features*

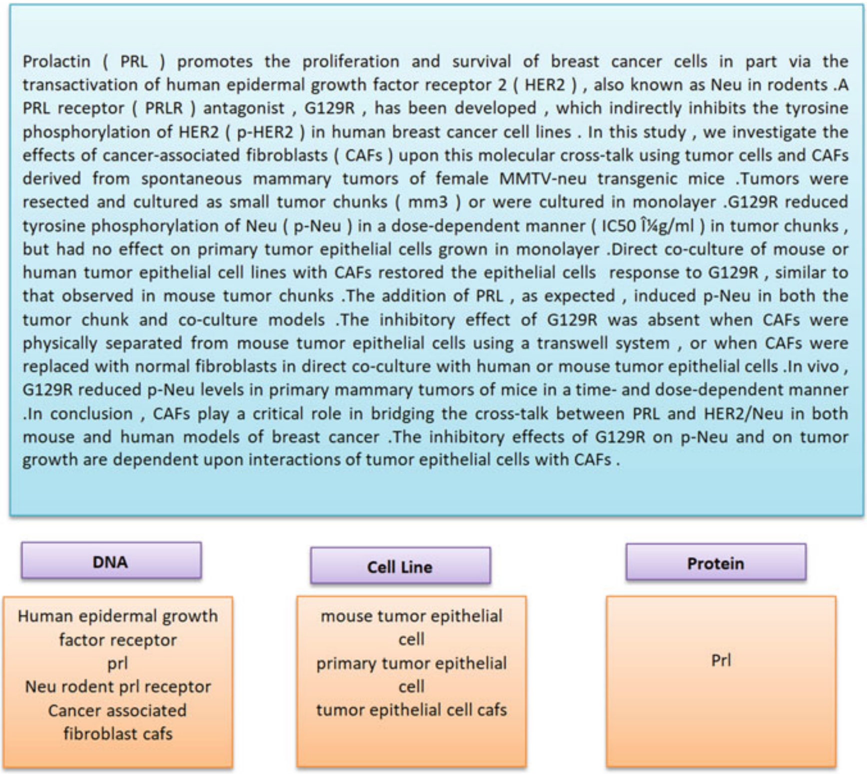
### Named Entity

The initial stage in biological text mining is to recognize named things. As a result, the subject has received a lot of attention in the bioinformatics community in recent years [8]. The job of named entity recognition (NER) in biological text mining is exceedingly significant and fundamental. To evaluate the performance of the entities, the JNLPBA [9] which is an open challenge task consisting of the five biomedical entity classes, namely from the GENIA [10] corpus, protein, cell line, cell type, DNA and RNA (Fig. 4).

The extraction of named entities and finding out their relationships (NERs) is one of the most fundamental and crucial tasks in biological data extraction. It entails extracting specific elements from text and categorizing them into predetermined categories. Protein, cell line, DNA, RNA and cell type are among the entities studied by JNLPBA. So here in our proposed model, one of the important categories of feature extraction is the identification of the biological words where we have used the JNLPBA in Python and the five major entities are recognized, i.e., name of the protein, the DNA, the RNA, cell line and the cell type.

### MeSH and List of Chemicals

The National Library of Medicine manages the Medical Subject Headings (MeSH) [11] vocabulary thesaurus, which indexes the MEDLINE documents. MeSH is made up of a set of description terms that are arranged in the form of the hierarchy (called MeSH trees) with more broad phrases closer to the root and more particular terms closer to the leaves [12]. Each MEDLINE text is manually annotated with a set of MeSH headings (typically 10–15), with three to five key headings representing the document's principal topics. MeSH is a hierarchical MeSH tree structure that has groups of "terms naming descriptors" that encompass the fields of medical, nursing, dentistry and veterinary medicine. As of 2022, MeSH has 30,193 main headings which increases every year and enhances the classification power of medical papers. So after giving the abstract as input from the dataset, the MeSH terms are extracted as one of the features in biomedical category (Fig. 5).



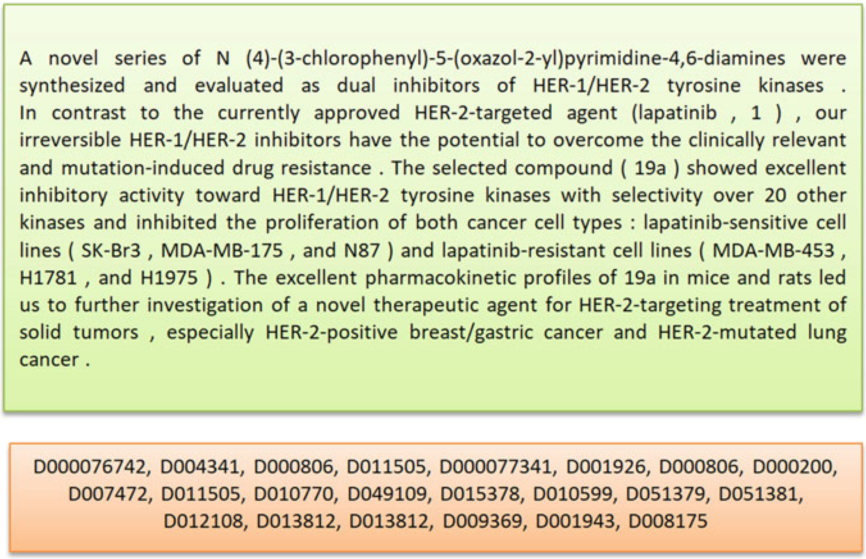
Prolactin ( PRL ) promotes the proliferation and survival of breast cancer cells in part via the transactivation of human epidermal growth factor receptor 2 ( HER2 ) , also known as Neu in rodents .A PRL receptor ( PRLR ) antagonist , G129R , has been developed , which indirectly inhibits the tyrosine phosphorylation of HER2 ( p-HER2 ) in human breast cancer cell lines . In this study , we investigate the effects of cancer-associated fibroblasts ( CAFs ) upon this molecular cross-talk using tumor cells and CAFs derived from spontaneous mammary tumors of female MMTV-neu transgenic mice .Tumors were resected and cultured as small tumor chunks ( mm3 ) or were cultured in monolayer .G129R reduced tyrosine phosphorylation of Neu ( p-Neu ) in a dose-dependent manner ( IC50 Î¼g/ml ) in tumor chunks , but had no effect on primary tumor epithelial cells grown in monolayer .Direct co-culture of mouse or human tumor epithelial cell lines with CAFs restored the epithelial cells response to G129R , similar to that observed in mouse tumor chunks .The addition of PRL , as expected , induced p-Neu in both the tumor chunk and co-culture models .The inhibitory effect of G129R was absent when CAFs were physically separated from mouse tumor epithelial cells using a transwell system , or when CAFs were replaced with normal fibroblasts in direct co-culture with human or mouse tumor epithelial cells .In vivo , G129R reduced p-Neu levels in primary mammary tumors of mice in a time- and dose-dependent manner .In conclusion , CAFs play a critical role in bridging the cross-talk between PRL and HER2/Neu in both mouse and human models of breast cancer .The inhibitory effects of G129R on p-Neu and on tumor growth are dependent upon interactions of tumor epithelial cells with CAFs .

**Fig. 4** Extraction of DNA, cell line and protein entities from an abstract from hallmark of cancer dataset

## 5.3 *Linguistic Features*

### Grammatical Relations

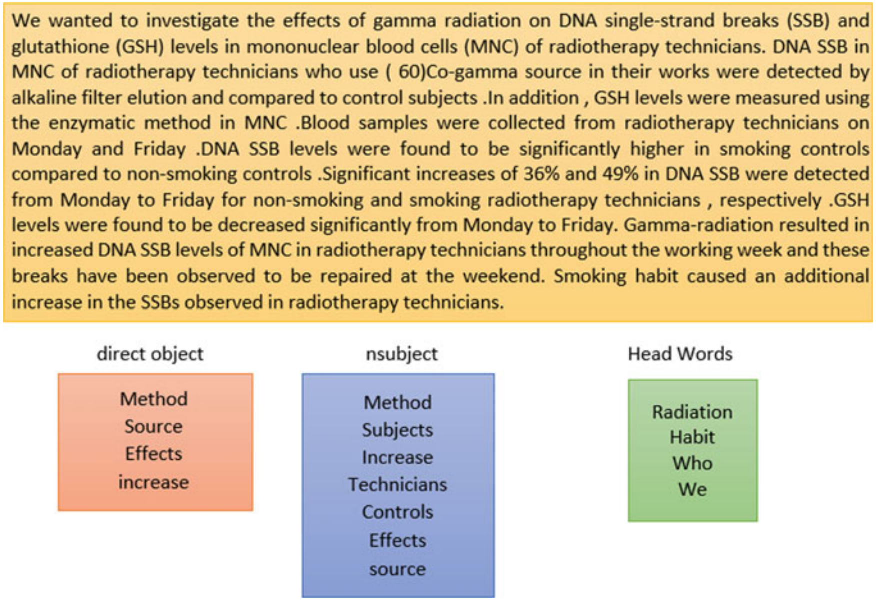
Every word in the language has a meaning so it is important to recognize and identify the words and their grammatical relations. The linguistic features give weightage to the soundness and clarity of logic and the accuracy of thinking. Hence, here the grammatical relations such as the direct object, the indirect object, the non-clausal subject and the head words are extracted as the feature from the abstract and the Fig. 6 shows an example of the words being recognized as grammatical entities from one of the abstract from Hallmark of Cancer dataset.



A novel series of N (4)-(3-chlorophenyl)-5-(oxazol-2-yl)pyrimidine-4,6-diamines were synthesized and evaluated as dual inhibitors of HER-1/HER-2 tyrosine kinases . In contrast to the currently approved HER-2-targeted agent (lapatinib , 1 ) , our irreversible HER-1/HER-2 inhibitors have the potential to overcome the clinically relevant and mutation-induced drug resistance . The selected compound ( 19a ) showed excellent inhibitory activity toward HER-1/HER-2 tyrosine kinases with selectivity over 20 other kinases and inhibited the proliferation of both cancer cell types : lapatinib-sensitive cell lines ( SK-Br3 , MDA-MB-175 , and N87 ) and lapatinib-resistant cell lines ( MDA-MB-453 , H1781 , and H1975 ) . The excellent pharmacokinetic profiles of 19a in mice and rats led us to further investigation of a novel therapeutic agent for HER-2-targeting treatment of solid tumors , especially HER-2-positive breast/gastric cancer and HER-2-mutated lung cancer .

D000076742, D004341, D000806, D011505, D000077341, D001926, D000806, D000200, D007472, D011505, D010770, D049109, D015378, D010599, D051379, D051381, D012108, D013812, D013812, D009369, D001943, D008175

**Fig. 5** Extraction of MeSH IDs from an abstract from hallmark of cancer dataset



We wanted to investigate the effects of gamma radiation on DNA single-strand breaks (SSB) and glutathione (GSH) levels in mononuclear blood cells (MNC) of radiotherapy technicians. DNA SSB in MNC of radiotherapy technicians who use ( 60)Co-gamma source in their works were detected by alkaline filter elution and compared to control subjects .In addition , GSH levels were measured using the enzymatic method in MNC .Blood samples were collected from radiotherapy technicians on Monday and Friday .DNA SSB levels were found to be significantly higher in smoking controls compared to non-smoking controls .Significant increases of 36% and 49% in DNA SSB were detected from Monday to Friday for non-smoking and smoking radiotherapy technicians , respectively .GSH levels were found to be decreased significantly from Monday to Friday. Gamma-radiation resulted in increased DNA SSB levels of MNC in radiotherapy technicians throughout the working week and these breaks have been observed to be repaired at the weekend. Smoking habit caused an additional increase in the SSBs observed in radiotherapy technicians.

| direct object | nsubject | Head Words |
|---|---|---|
| Method<br>Source<br>Effects<br>increase | Method<br>Subjects<br>Increase<br>Technicians<br>Controls<br>Effects<br>source | Radiation<br>Habit<br>Who<br>We |

**Fig. 6** Linguistics feature extraction from a sample abstract from hallmark of cancer dataset

## 6 Experiments and Results

This section describes in detail the various approaches applied to the dataset to classify the abstracts of the PubMed articles.

According to the paper [13] as cited in the reference also it has achieved the accuracy as mentioned in the column and we tried to improve the accuracy using different techniques like.

### Individual feature

Here, we have used individual features like first we provided only the statistical features and as described in the previous section the noun bigrams and lemmatized bag of words forms the only two features which were given to the classifier SVM to classify the abstract. Then, we tried the same method with linguistic and then with only the biomedical features but according to our observation, the individual features play less important role to improve the accuracy, then Approach-2 was tested on the dataset.

### Hybrid Technique

In the second approach, we experimented with hybridization of the features where initially we combined the biomedical features with the statistical features (approach 2a) same with statistical features combined with the linguistic features (approach 2b) and finally with biomedical features in combination with the linguistic features, and according to the experiment, we found out that hybridization of biomedical features with linguistic features is giving the best accuracy and has also increased the performance of the model stated by Baker [13].

The average accuracy has also increased from 94.92 to 98.70%. Even for individual classes like: Sustaining proliferative signaling the accuracy has increased from 88.90 to 98.64%, evading growth suppressors from 91.70 to 98.37%, for the class resisting cell death, the accuracy has improved from 95.50 to 98.37%, class 4: Enabling replicative immortality, there is an increase of 1.78% (from 97.40 to 99.18%) for the class inducing angiogenesis, accuracy achieved is 99.05% and same for the activating invasion and metastasis. For the class genomic instability and mutation, accuracy has increased to 98.64% and for the class tumor promoting inflammation to 98.37% and for cellular energetics, it is 98.64% and for the last class, avoiding immune destruction accuracy achieved is 98.37% using this approach of hybridization of biomedical and linguistic features (Table 1).

**Table 1** Comparison of the accuracy measure using our approach (individual feature and hybrid technique) and the paper [13]

| S. No | Hallmark | Accuracy [13] | Individual feature | | | Hybrid technique | | |
|---|---|---|---|---|---|---|---|---|
| | | | Statistical | Biomedical | Linguistic | Approach-2a | Approach-2b | Approach-2c |
| 1 | Sustaining proliferative signaling | 88.90 | 81.57 | 80.27 | 81.57 | 79.86 | 81.57 | **98.64** |
| 2 | Evading growth suppressors | 91.70 | 88.86 | 88.86 | 88.86 | 89.32 | 89.87 | **98.37** |
| 3 | Resisting cell death | 95.50 | 82.18 | 82.18 | 82.18 | 82.97 | 82.18 | **98.37** |
| 4 | Enabling replicative immortality | 97.40 | 94.31 | 94.30 | 94.33 | 95.54 | 94.93 | **99.18** |
| 5 | Inducing angiogenesis | 98.40 | 93.72 | 93.72 | 93.72 | 94.50 | 93.72 | **99.05** |
| 6 | Activating invasion and metastasis | 94.00 | 87.85 | 87.80 | 87.85 | 89.72 | 87.85 | **99.05** |
| 7 | Genomic instability and mutation | 91.70 | 87.65 | 87.60 | 87.65 | 87.29 | 86.23 | **98.64** |
| 8 | Tumor promoting inflammation | 95.90 | 91.09 | 91.09 | 91.09 | **98.67** | 91.09 | **98.37** |
| 9 | Cellular energetics | 99.60 | 97.36 | 94.86 | 97.36 | 95.54 | 97.36 | 98.37 |
| 10 | Avoiding immune destruction | 96.10 | 95.34 | 95.27 | 95.34 | 94.59 | 94.53 | **98.91** |
| | | 94.92 | 89.99 | 89.60 | 90.00 | 90.80 | 89.93 | 98.70 |

## 7 Conclusion

With the increasing number of daily medical research, it has become very important to develop efficient methods for automatically classifying text data. Binary classification methods are easy to classify the problems with 2 classes but the multi-label classification problems are hard to tackle and are really challenging when the number of classes increases. Our evaluation shows the hybridization of different features when provided as an input to the supervised learning algorithm SVM leads to the improvement of the results specially in case of biomedical and linguistic features. Support vector machines provide great results in case of the multi-label classification problem with the benchmark dataset of Hallmark of Cancer. In the future work, research could focus more on using the biomedical features along with different text representation techniques to reduce the sparsity of the vectors and also implementation of the CNN model.

## References

1. McEntyre J, Lipman D (2001) PubMed: bridging the information gap. Can Med Assoc J 164:1317–1319
2. Pubmed <http://www.ncbi.nlm.nih.gov/pubmed/>.
3. Guansong P, Shengyi J (2012) Technology research of automatic text classification. Inf Studies Theory Appl 35(2):123–128
4. Ozturkmenoglu O, Alpkocak A (2012) Comparison of different lemmatization approaches for information retrieval on Turkish text collection. Int Symp Innov Intell Syst Appl 2012:1–5. https://doi.org/10.1109/INISTA.2012.6246934
5. Kalcheva N, Karova M, Penev I (2020) Comparison of the accuracy of SVM kemel functions in text classification. Int Conf Biomed Innov Appl (BIA) 2020:141–145. https://doi.org/10.1109/BIA50171.2020.9244278
6. Modi, n.d.) Nareshpalsingh, J. M., & Modi, H. N. (n.d.). Multi-label Classification Meth-ods: A Comparative Study. 04(12), 8.
7. Lemma (morphology) (2022) In Wikipedia. https://en.wikipedia.org/w/index.php?title=Lemma_(morphology)&oldid=1067206748
8. Cheng W, Hüllermeier E (2009) Combining instance-based learning and logistic regression for multilabel classification. Mach Learn 76(2):211–225. https://doi.org/10.1007/s10994-009-5127-5
9. Kim J-D, Ohta T, Tsuruoka Y, Tateisi Y, Collier N (2004) Introduction to the bio-entity recognition task at JNLPBA. In: Proceedings of the international joint workshop on natural language processing in biomedicine and its applications, pp 70–75
10. Kim J-D, Ohta T, Tateisi Y, Tsujii J (2003) GENIA corpus—a semantically annotated corpus for bio-textmining. Bioinformatics 19(suppl_1):i180–i182. https://doi.org/10.1093/bioinformatics/btg1023
11. Home—MeSH—NCBI. (n.d.). Retrieved 27 Feb 2022. https://www.ncbi.nlm.nih.gov/mesh/
12. Jing Z, Yuxuan S, Shengwen P, Xuhui L, Hiroshi M, Shanfeng Z (2015) MeSHSim: an R/Bioconductor package for measuring semantic similarity over MeSH headings and MEDLINE documents. In: 2015 34th Chinese control conference (CCC), pp 8535–8539. https://doi.org/10.1109/ChiCC.2015.7260989
13. Baker S, Korhonen A, Pyysalo S (n.d.) Cancer hallmark text classification using convolutional neural networks 26

14. Tsoumakas G, Katakis I (2009) Multi-label classification: an overview. Int J Data Warehouse Min 3:1–13. https://doi.org/10.4018/jdwm.2007070101
15. Wang Z, Sun X, Zhang D (2006) An optimal text categorization algorithm based on SVM. In: 2006 International conference on communications, circuits and systems, vol 3. Pp 2137–2140. https://doi.org/10.1109/ICCCAS.2006.284921

# COVID-19 Prediction from CT and X-Ray Scan Images: A Review

**Prashant Kumar and Jagdeep Kaur**

**Abstract** The epidemiological data containing the number of individuals who are infected earlier and the total population is employed in the predictive models of coronavirus 2019. Different parameters namely latency period and the possibility of curing from the disease are utilized for predicting the infection trends. But, these models are inefficient to reflect diverse socio-static and economic factors due to which the course of the virus is affected. The COVID-19 prediction techniques are based on pre-processing, feature extraction and classification. The various machine learning and segmentation algorithms are proposed so far for the coronavirus disease 2019 prediction from the X-Ray images. This paper studies the comparison of deep learning techniques, machine learning techniques and ensemble techniques, various popular schemes are reviewed in terms of their results as well as their limitations.

**Keywords** COVID-19 · X-Ray · Segmentation · Feature Extraction · Classification

## 1 Introduction

COVID-19 is transmitted at an extensive level all over the globe. WHO declared this disease as a pandemic. The number of individuals, who suffered from COVID-19, is more than 16.562 million in which many people are died and some are recovered. The actual number of infected cases is more than the number of confirmed patients. To recognize Corona disease, RT-PCR is available as a standard mode. However, this procedure consumed a lot of time and the primary virus concentration is lower and simple FNs (False Negative) are occurred [1]. Thus, CT (Computed Tomography) and X-rays images are effective to perform the auxiliary diagnosis procedure. These tools are useful to diagnosing and recognizing the growth of disease. In addition, other significant procedures are to evaluate the lung infections from CT (Computed Tomography) and X-ray images, testing, to observe the isolation and to cure this

---

P. Kumar (✉) · J. Kaur
Dr B R Ambedkar National Institute of Technology, Grand Trunk Road, Barnala–Amritsar Bypass Rd, Jalandhar, Punjab 144011, India
e-mail: prashantkumar.78698@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_18

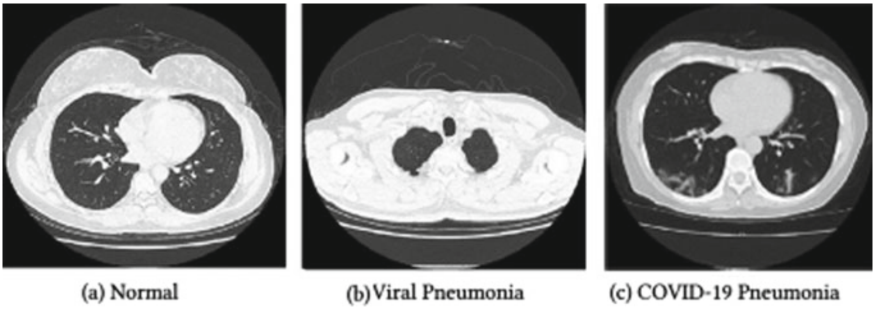
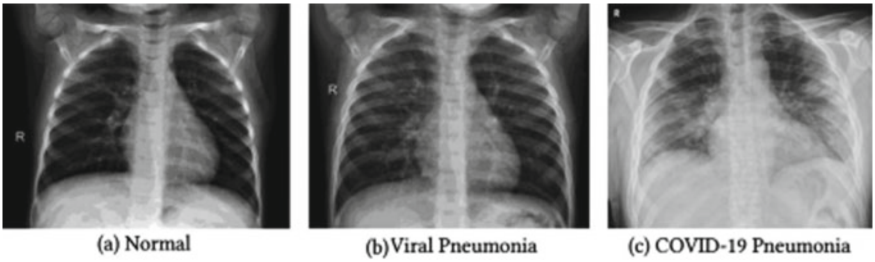

(a) Normal (b) Viral Pneumonia (c) COVID-19 Pneumonia

**Fig. 1** CT scan images



(a) Normal (b) Viral Pneumonia (c) COVID-19 Pneumonia

**Fig. 2** CXR images

disease. CT scans aid in detecting mutated COVID-19 in comparison with RT-PCR due to its potential of detecting FN (false negative). The approach of computing the treatment effect becomes easy on the basis of CT scans (Figs. 1 and 2).

Chest radiography imaging such as X-ray or CT, that is a custom method to diagnose the pneumonia, can be easily applied [2]. Corona virus disease 2019 can be detected quickly and at higher sensitivity using this method. The CXR (Chest X-ray) images are capable of representing the visual indexes related to COVID-19. Different studies prove radiography as a feasible detection tool to detect COVID-19. Up till now, there is not any study conducted on the potential of AI (artificial intelligence) to detect corona disease in automatic manner from X-ray or chest CT images as the public images are not available in adequate amount from COVID-19 patients. At present, a small dataset of CXR images of this disease is generated through some researchers for training AI models so that corona disease can be diagnosed automatically. The extraction of these images was done from the academic publications reporting i.e., the outcomes of corona disease X-ray and CT imaging [3].

### 1.1 *COVID-19 Identification of Lung Infections with CT Imaging and Chest X-Ray (CXR)*

DIP (digital image processing) technology has attained a lot of attention in the medical field, such as to segment an organ and to enhance and restore an image for assisting in subsequent medical diagnoses, as the computer technology is progressing rapidly. A NN (neural network) is introduced via AI to simulate the human brain in order to analyses and learn the way of promoting ML (machine learning) [4]. NN (neural network) is planned on the basis of human brain to accomplish an analysis and processing. Various sectors such as computer graphics, image processing, computer vision, and computer-aided design make the deployment of integrated AI technology, deep mining of big data, and optimization of computational models. The scientific techniques are generated by remedial effect evaluation and proceedings which are becoming major scientific problems and key technologies of cutting-edge medical imaging. There is necessity to tackle this issue in the field of medical image analysis.

The particular process of identifying the COVID-19 of Lung Infections with CT Imaging is represented in the given Fig. 3. This process is consisted of various phases which are defined as:

i. **Pre-processing**. The input CT (computed tomography) is pre-processed for extracting efficient lung regions. It is essential to extract the attributes from the CXR images for detecting the infection of Corona virus Disease 2019 on the basis of automatic classifier [5]. This phase employs a process of binarization of an image for lessening the amount of information subjected to accomplish the further analysis and boosting up its speed. In this process, the colored image is transformed into an image having only black and white pixels. For this, a threshold is fixed as it is a basic rule.

ii. **Candidate Region Segmentation**. In the process of detecting the lung infection, the physicians majorly focus on locating the affected region and extracting the outline in accurate manner on the basis of local symptoms later on. The significant attributes are the normal region and the junction region for differentiating the normal tissue from the affected one. Several studies make the deployment of ML-based (Machine Learning) image segmentation in chest images having corona disease. The medical image is segmented using U-Net. This network is adopted by various researchers for segmenting the COVID-19 lung image.

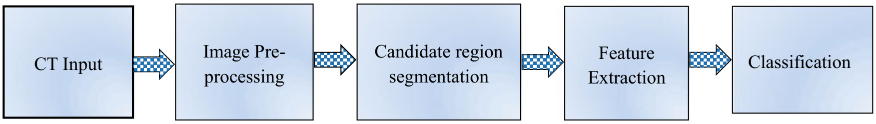


**Fig. 3** Process of identification of COVID-19 of lung infections with CT imaging

U-Net is a fully CNN (Convolutional neural network) with a U-shaped structure having symmetrical encoding and decoding signal paths. The shortcuts are employed to connect the similar layer. A U-shaped SGD + Momentum structure is generated on the basis of four down-sampling and four up-sampling [6].

iii. **Feature Extraction**. In this procedure, a portion of data is transformed into only a relevant group of bits. The technique of extracting the attributes is adopted in research of classifying an image for diminishing the number of pixels contained in the classification algorithm for learning [7].The major method of extracting features is texture features for the X-ray images due to their potential for evaluating and localizing the groups of pixels of the same intensity. Their testing is already done in the field of classifying X-ray images. GLCM algorithm (Gray Level Co-Occurrence Matrix) is deployed for extracting the attributes from the chest X-rays. This matrix is able to store the number of times a pixel has certain gray level intensity and occurred with another pixel in the image. For this, a particular rule or configuration called offset is implemented in which a pattern of horizontal comparison at 0°, vertical at 90° or diagonally at 45°, or 135° is considered. The fundamental goal of this procedure is to acquire the spatial associations of the pixels related to diverse intensities of the image and for storing them in the GLCM. Moreover, the extraction of six statistical attributes is done from the GLCM (Gray Level Co-Occurrence Matrix) of every image, and this is in correspondence with six texture operators.

iv. **Classification**. When the attributes are extracted, the major task is to classify the corona disease from normal and other pneumonia. To classify the images, the moving of network is required into the fully connected layer for which the images are flattered and simplified. A complex set of NN (Neural Network) connections are capable only of processing the linear data. In case, the data is not labelled, auto encoders are deployed to implement the unsupervised ML (Machine Learning) algorithms for assembling the data in a low dimension space performing calculations [8]. Thereafter, the existing data is up-sampled by reconstructing the additional layers.

## *1.2 Artificial Intelligence in Medical Image Analysis*

Artificial Intelligence is related to computer science field which assists machines in executing the human intelligence tasks [9]. As the Artificial Intelligence and IoT is developed, continues change can be seen in medical equipment that offers various possibilities in medical radiology. The Machine Learning methods are a subset of Artificial Intelligence which provides learning potential to computer systems and assists in executing the tasks with the data in automatic manner without manual programming. DL is a subset of ML which deals in the simulation of the neurons of the human brain. Machine Learning focuses on deploying the Deep Learning as a crucial subject with its technology in order to classify, recognize and identify the images or videos. Diverse DL and ML useful for identifying Corona virus Disease 2019 are explained as:

i. **Convolutional neural network (CNN)**. CNNs are generally implemented in image processing problems at which a computer recognizes an object in a picture. Its modelling is sufficient to process and classify images. The data is considered as spatial. Different from normal NN (Neural Network), neurons of this algorithm are not linked to others, and eventually, the output layers focus only on selecting the nearest neurons with the similar weights. This algorithm retains the spatial aspect of the dataset. Hence, a filtering process is applied on it to simplify the complicated images which undergone from superior processing [10]. This algorithm consists of multiple layers, with many different layers called the convolutional layer, the pooling layer, and the fully connected layer. The initial layer also has a ReLU which leads to activate functions and guarantees non-linearity due to the progressing of data via every network layer. In the absence of this unit, the data offered at the input layer may misplace the dimension which is essential in the network. The last layer is utilized to perform classification on the dataset. A filter is placed on an array of image pixels and a complex feature map is also developed to operate this algorithm. The analogy is like looking at an image by a window that allows one to see specific features within the image. It is also called typical two-dimensional CNN.
ii. **Support Vector Machine (SVM)**. It is exploited for investigating the hyperplane. The decision boundary is discovered along with hyperplane using this algorithm and a lower error rate is acquired. The hyperplane in a SVM is expressed as:

$$y = w^t \cdot x + b \tag{1}$$

In which, the vector of features is represented with $x$, $w$ illustrates a vector of normalized direction to the hyperplane, and a form of threshold is denoted by $b$ [11]. SVM is an effectual algorithm using which dataset is projected into higher dimension based on kernels. A common kernel function $K$ is present as a linear such as dot product or non-linear in form of RBF (Radial basis function) kernel which is expressed as:

$$\|y - Hx\|_2 <\in K(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2} \tag{2}$$

In this, a parameter for rbf kernel is defined with $\gamma$ and $\|x_1 - x_2\|^2$ represents the Euclidean distance.

iii. **Random Forest (RF)**. RF (Random Forest) technique is implemented for constructing a forest of individual DTs (decision trees) which results in composing an ensemble [12]. The majority votes are employed to predict the data via every tree. Consequently, the class with the majority vote is considered as the final prediction. Therefore, the effective prediction is obtained on the data with the integration of individual uncorrelated systems. RF is capable of maintaining the correlation among trees as minimal as possible. For this purpose, Bagging is exploited. Due to the sensitivity of DTs to training data, RF concentrates in building the trees on data that is sampled at random with replacement. As a result, various trees are produced with minimum correlation amongst them. In addition, Random Forest algorithm is efficient to sustain Feature Randomness because of the potential of every tree of making its decision on the basis of random subset of features instead of picking each attribute from the dataset. This renders a great variation in the factors at which every tree is set up. Therefore, bagging and feature randomness assisted this algorithm in training with random samples of a dataset as well as with different subsets of features to achieve prediction. Moreover, the hyper parameter tuning of a sufficiently large parameter set is useful to prevent it from over fitting on the data [13].

## 2 Review Methodology

This section presents the state-of-the-art review layout, a step-by-step method for the literature discussed in the previous sections. This research focuses on categorizing the current literature on COVID-19 detection assessing the current trends. This evaluation finds relevant research articles from reputable electronic databases and the top conferences in the field. After then, inclusion and exclusion criteria were used to reduce the number of papers that were considered. Following that, final research studies were chosen based on a variety of variables. The information given here is the product of a thorough investigation. For this review study, various electronic database sources were investigated; some of the popular electronic databases used in this search like google scholar, Elsevier, Science direct etc. Using the inclusion criterion, which mainly depends on the techniques, the relevant work of COVID-19 detection algorithms is retrieved from the enormous collection of data given by search engines. The data shows that journals account for most of the work in this study (51%), with conferences accounting for 40% of the work and book chapters accounting for 9%. In addition, the data depicts a year-by-year study of work relevant to diabetic retinopathy detection. The major data is available on the google scholar as compared to Elsevier and Science direct. The google scholar has 60% data, Elsevier has approx. 10% and Science direct has approx. 30 data on diabetic retinopathy detection. The data division is shown in Fig. 4.

As shown in Fig. 4, the percentage of data sharing is shown in figure approx. 60% data is available on Google scholar, 30% is available on science direct and very less amount of data that is 10% is available on Elsevier.

As shown in Fig. 5, the data is available which is available through conference, journal and books. The conference has approx 51% of total data, 40% data has available through journal and 9% is available through books.

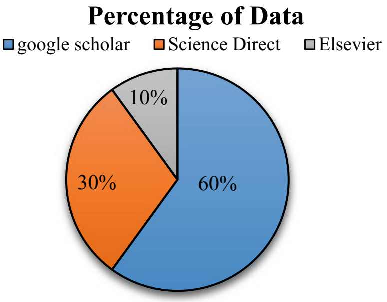


**Fig. 4** Percentage of data sharing

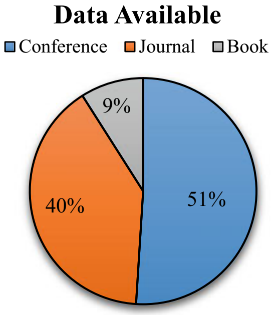


**Fig. 5** Data available

# 3 Literature Review

## *3.1 COVID-19 Identification of Lung Infection Using Deep Learning*

Visvanathan et al. emphasized on implementing a CNN (convolutional neural network) algorithm for recognizing the Coronavirus Disease 2019 and learning from multiple data sources allocated for only a few classes [14]. Three public datasets were employed to train the presented algorithm and compute its efficiency for segmenting the lung field, lobes and corona disease infected areas. The evaluation depicted that the presented algorithm offered an average error around 4.5% while identifying the infection per lobe in 8CT (Computed Tomography) scans of lung.

Abdulla et al. suggested a DCNN (Deep convolutional neural networks) called SAARS Net with the purpose of detecting coronavirus diseases 2019 cases from CXR images [15]. In addition, the potential of the suggested model, for differentiating the corona class from normal and infected classes and to attain deeper insights into critical factors of Coronavirus Disease 2019 cases, was analyzed under 3 scenarios. This model was quantified on the confusion parameters. The results demonstrated that the suggested model provided the positive predictive value up to 100% and generated optimal outcomes to detect COVID-19 and abnormal classes.

Islam et al. introduced a simple CNN (convolution neural network) based system with the help of lung CT (Computed Tomography) images to recognize the Coronavirus Disease 2019 [16]. Moreover, to achieve this, LeNet-5 CNN model was deployed. A technique of augmenting the data was put forward and attained various Computed Tomography images of covid positive and normal classes. The accuracy acquired from the introduced system was computed 86.061%, $f1$ score was 87.1%, the precision was 85.02%, recall was 89%, and AUROC was calculated 0.86 while detecting corona disease.

Abdar et al. projected a DL-based CNN in order to distinguish the corona disease 2019 positive patients from normal subjects via chest CT (Computed Tomography) images [17]. The training, testing and computation of projected technique was done on data in which 10,979 chest Computed Tomography images of patients and 150 normal subjects were included. The outcomes indicated that the projected technique had provided 92% precision, 90% sensitivity, 91% specificity, 0.91 $F1$-Score and 90% accuracy.

Jain et al. designed a CNN (Convolutional Neural Network) algorithm for diagnosing Corona virus Disease 2019 disease from CXR (chest X-ray) [18]. The multiple convolutional layers of this algorithm assisted in extracting the attributes of corona disease from CXR images. The reliability of this algorithm was proved and it had generated the results at instant with lower computational effort for determining whether Covid-19 induced pneumonia was occurred or not. The results exhibited that the designed algorithm was useful for radiologists for making decisions on X-ray films so that Corona disease was detected.

Satyabrata et al. used the support vector regression (SVR) and Long Short Term Memory (LSTM) techniques to simulate the behavior of the COVID pandemic [19]. The SVR technique for time series forecasting is a helpful tool for estimating objective value functions. It is used to forecast the value of a continuous variable with a known value. For the prediction of COVID-19 behavior, the LSTM technique is a better forecasting approach. LSTM addresses the disadvantage of RNNs (Recurrent Neural Network). The performance of LSTM is 95.46%, which is significant enough to proceed with pandemic behavior predicting (Table 1).

### 3.2 *COVID-19 Identification of Lung Infection Using Machine Learning*

Panetta et al. established a new shape-dependent Fibonacci-p patterns-based feature descriptor on the basis of ML (machine learning) technique [20]. Diverse metrics namely accuracy, precision, and f1-score were utilized to analyze the established technique on Kaggle and COVIDGR datasets. The results revealed that the established technique yielded the efficiency of 100% for distinguishing the normal and COVID-19 from CT images of chest of initial dataset. Moreover, this technique provided recall around 72.65 ± 6.83 and specificity of 77.72 ± 8.06 on the subsequent dataset.

**Table 1** Comparison of deep learning techniques

| Author | Year | Technique used | Findings | Limitations |
|---|---|---|---|---|
| Visvanathan et al. | 2021 | Convolutional neural network | The evaluation depicted that the presented algorithm offered an average error around 4.5% while identifying the infection per lobe | This algorithm provided mower efficiency for segmenting the infected areas |
| Abdulla et al. | 2020 | DCNN (deep convolutional neural network) called SAARS Net | The results demonstrated that the suggested model provided the positive predictive value up to 100% and generated optimal outcomes to detect COVID-19 and abnormal classes | This model was not capable of detecting the corona disease at initial phases. Moreover, the problem related to the imbalanced images in diverse classes was occurred due to which results of chest X-Rays were affected |
| Islam et al. | 2020 | Simple convolution neural network (CNN) | The accuracy acquired from the introduced system was computed 86.061%, $f1$ score was 87.1%, the precision was 85.02%, recall was 89%, and AUROC was calculated 0.86 while detecting corona disease | This introduced system was not suitable for a larger dataset for dealing with this pandemic |
| Jain et al. | 2021 | CNN (Convolutional Neural Network) | The reliability of this algorithm was proved and it had generated the results at instant with lower computational effort for determining whether Covid-19 induced pneumonia was occurred or not | The designed algorithm offered lower efficacy in case of large repositories of CXR (chest X-ray) films and CT (Computed Tomography) scan images |

(continued)

**Table 1** (continued)

| Author | Year | Technique used | Findings | Limitations |
|---|---|---|---|---|
| Abdar et al. | 2020 | DL-based CNN algorithm | The outcomes indicated that the projected technique had provided 92% precision, 90% sensitivity, 91% specificity, 0.91 $F1$-Score and 90% accuracy. The results were made more reliable and generalized using the infected areas | The data was available in inadequate amount for maximizing the generalization of the system for practical applications |

Lucas et al. intended a TL (transfer learning) method on a pre-trained VGG19 with the objective of constructing a DCNN (deep convolutional neural network) for detecting 4 possible conditions such as normal, bacteria, abnormal and Corona disease positive [21]. The investigated patterns had reliability for images from a given class of disease and led to offer explanations of process of DL (Deep Learning) model for making decisions regarding classification. These patterns were utilized as biomarkers for a given disease in CXR images. The intended method was able to detect the Coronavirus Disease 2019 efficiently and obtained the accuracy of 100%.

Zebari et al. developed a fusion mechanism on the basis of ML (machine learning) system for with regard to 3 significant texture attributes such as LBP (Local Directional Patterns), FD (Fractal Dimension) and GLCM (Grey Level Co-occurrence Matrix) [22]. This mechanism focused on recognizing the chest CT (Computed Tomography) scan image as normal or infected with Coronavirus Disease 2019. A public dataset, consisted of 300 CT images, was applied for validating the developed mechanism. The experimental outcomes confirmed that the developed mechanism provided the accuracy of 96.91% and the accuracy of 89.87% for LBP attribute, 87.84% for FD, and 90.98% for GLCM.

Arora et al. recommended an automatic technique recognized as XG Boost in order to diagnose novel COVID-19 through CXR images [23]. This technique was planned on the basis of convolutions layered mode. The recommended was adaptable for attaining the accuracy of 90%. The results of this technique were compared with the traditional models. The applicability of the recommended technique was proved for diagnosing the corona disease.

Absar et al. introduced a ML technique recognized as SVM to diagnose Coronavirus Disease 2019 from CXR images [24]. More images were attained using image processing tool known as RGB and Squeeze Net models so that the given data set was diagnosed. The accuracy of the introduced algorithm was computed 98.8% for detecting patients, who suffered from COVID-19, from CXR images. The introduced

CAT (computer-aided detection tool) was efficient to mitigate the transmission of infectious diseases in society by means of a faster patient screening process (Table 2).

### 3.3 COVID-19 Identification of Lung Infection Using Ensemble Techniques

Sanagavarapu et al. devised an ensemble technique for tackling the issue of class imbalance when the positive COVID-19 chest CT (Computed Tomography) scans were recognized [25]. The CLAHE technique was implemented to normalize the images in order to enhance their attributes. After that, the dataset was divided into equally balanced portions and the Level-0 Res Net models employed the images to train the technique. The devised technique was able to offer and accuracy up to 87.23% with CLAHE enhanced images and GradCAM algorithm was useful to display the learning of the trained models in visual form.

Foysal et al. suggested an ensemble DCNN technique to diagnose Corona virus Disease 2019 cases from the chest CT images [26]. Several hyper-parameters were considered with the purpose of planning, training and validating the DCNN models. The predictive values of these algorithms were integrated. Consequently, the accuracy was enhanced and the error rate was diminished. An ensemble hard voting rules were adopted for classifying the COVID cases from the normal ones. The results of experimentation depicted that the suggested technique yielded the accuracy around 96% and sensitivity up to 97%.

Afshar et al. presented a fully-automated model on the basis of Capsule Networks called CT-CAPS, for distinguishing Coronavirus Disease 2019 from healthy and CAP cases with the help of chest CT (Computed Tomography) scans [27]. An enormous amount of medical data related to gender, age, weight and symptoms of patient was exploited to enhance the presented model. RF (Random Forest) algorithm was adopted in this hybrid model in which medical data and CT scans were applied for classifying corona disease and healthy. The outcomes of experiment validated that the hybrid model assisted in enhancing the efficacy of the presented approach and offering the precision up to 90.8%, sensitivity around 94.5% and specificity up to 86.0%.

Nneji et al. formulated a novel joint framework in order to recognize the Corona virus Disease 2019 accurately in which an enhanced super-resolution GAN was put together with a filter bank of reducing noise of WT-CNN on CXR and CT images of corona disease [28]. The formulated framework had capacity for eliminating the noise and gradient problem. The formulated framework offered 98.8% accuracy, 98.2% $F1$-score and 98.9% precision in case of CXR dataset and 97.8% accuracy, 96.1% $F1$-score and 98% precision for CT images dataset in comparison with traditional techniques.

**Table 2** Comparison of machine learning techniques

| Author | Year | Technique used | Findings | Limitations |
|---|---|---|---|---|
| Panetta et al. | 2021 | Novel shape-dependent Fibonacci-p patterns-based feature descriptor using a ML | The established technique yielded the efficiency of 100% for distinguishing the normal and COVID-19 from CT images Moreover, this technique provided recall around 72.65 ± 6.83 and specificity of 77.72 ± 8.06 on the subsequent dataset | The established technique was not useful for analyzing long-term impact of Corona virus Disease 2019 from the patients who were recovered from the disease |
| Lucas et al. | 2020 | Transfer learning technique | The intended method was able to detect the Corona virus Disease 2019 efficiently and obtained the accuracy of 100% | The experiments were restricted due to this method as it was planned on the basis of pre-trained VGG-19 |
| Zebari et al. | 2020 | A fusion scheme based on a ML technique | The experimental outcomes confirmed that the developed mechanism provided the accuracy of 96.91% and the accuracy of 89.87% for LBP attribute, 87.84% for FD, and 90.98% for GLCM | The efficiency of this mechanism to diagnose the corona disease was mitigated in the presence of geometric features |
| Arora et al. | 2021 | XGBoost | The recommended was adaptable for attaining the accuracy of 90%. And applicability of the recommended technique was proved for diagnosing the corona disease | This technique provided lower precision due to the finite amount of COVID-19 data |
| Absar et al. | 2022 | SVM (support vector machine) | The introduced CAT (computer-aided detection tool) was efficient to mitigate the transmission of infectious diseases in society by means of a faster patient screening process and yielded accuracy of 98.8% | The introduced algorithm yielded lower scalability due to the imbalanced dataset |

Mandal et al. formulated a novel zero-trust network-based access control policy By specifically restricting incoming network traffic to substantiate MAC spoofing tactics on the software-defined network (SDN) architecture of cloud computing [29]. The suggested technique focuses on a wireless network (IEEE 802.11) unaffected by physical attributes or location. The adoption of the threshold stamping technique reduces the false-positive rate (FPR) by allowing a legitimate user's traffic to be rectified before being classified to the attacker (FPR). In a worst-case situation, the highly skilled attacker's presence prevents the thresholds from decreasing, causing the TCP packet to be rejected by the algorithm and preventing the faked user from gaining access to the SDN framework of cloud architecture. Analyzing data and deleting a fake user are time-consuming tasks that may be time-consuming user time time-consuming tasks that may be easier with a little more investigation.

Gour et al. developed a novel stacked CNN algorithm to diagnose Corona virus Disease 2019 disease automatically from the CXR and Computed Tomography images [30]. VGG19 and the Xception algorithms offered various sub-models to train this algorithm. The discrimination efficacy of dissimilar sub-models of convolutional neural network was integrated with the efficacy for detecting Coronavirus Disease 2019 from the radiological images. The sensitivity acquired from the developed algorithm was counted 97.62% while classifying the X-ray images as COVID-19, healthy and Pneumonia Classes and performed better in contrast to the conventional techniques (Table 3).

## 4 Conclusion

In recent times, the major concern regarding this disease is that there is not any symptom shown in the patient who suffers from corona disease. These kinds of affected patients cause the transmission of disease and infected other people all over the world. Various reasons are present which transmit the 2019-novel corona virus rapidly. The dominant reason is the lack of information transparency in the premature stage of the epidemic outbreak. The corona virus prediction has diverse stages in which data is pre-processed, features are extracted and classification is done. It is analyzed that major techniques are unable to achieve optimal level of accuracy. The technique needs to propose which can achieve optimal level of accuracy for the COVID-19 prediction.

**Table 3** Comparison of ensemble techniques

| Author | Year | Technique used | Findings | Limitations |
|---|---|---|---|---|
| Sanagavarapu et al. | 2021 | An ensemble method | The devised technique was able to offer and accuracy up to 87.23% with CLAHE enhanced images and Grad CAM algorithm was useful to display the learning of the trained models in visual form | The devised technique was ineffective for diagnosing and curing the patients for their recovery at initial phases |
| Foysal et al. | 2021 | An ensemble DCNN | The results of experimentation depicted that the suggested technique yielded the accuracy around 96% and sensitivity up to 97% | The suggested technique provided poor performance on other diverse data |
| Afshar et al. | 2021 | ELM, Extreme learning machine with a hybrid CNN-PCA based feature extraction, -ELM with CLAHE | The sensitivity acquired was counted 97.62% while classifying the X-ray images as COVID-19, healthy and Pneumonia Classes | The developed algorithm was not able for classifying the Chest X-Ray images into diverse categories of corona disease cases |
| Nneji et al. | 2021 | Novel joint framework | The formulated framework offered 98.8% accuracy, 98.2% $F1$-score and 98.9% precision in case of CXR dataset and 97.8% accuracy, 96.1% $F1$-score and 98% precision for CT images dataset | This framework was not performed well in some scenarios of poor-quality images |
| Gour et al. | 2022 | New stacked convolutional neural network model | The outcomes of experiment validated that the hybrid model assisted in enhancing the efficacy of the presented approach and offering the precision up to 90.8%, sensitivity around 94.5% and specificity up to 86.0% | The efficiency of the presented model was lessened on large sized and multi-center datasets |

## References

1. Hastuti ET, Bustamam A, Anki P, Amalia R, Salma A (2020) Performance of true transfer learning using CNN densenet121 for COVID-19 detection from chest X-ray images. In: IEEE International conference on health, instrumentation & measurement, and natural sciences (InHeNce), vol 56(2), pp 367–374
2. Dutta P, Roy T, Anjum N (2021) COVID-19 Detection using transfer learning with convolutional neural network. In: 2nd International conference on robotics, electrical and signal processing techniques (ICREST), vol 2(12), pp 599–606
3. Shamsi A, Asgharnezhad H, Jokandan SS, Khosravi A, Kebria PM, Nahavandi D, Nahavandi S, Srinivasan D (2021) An uncertainty-aware transfer learning-based framework for COVID-19 diagnosis. IEEE Trans Neural Netw Learn Syst 11(6):515–523
4. Abbas A, Abdelsamea MM, Gaber MM (2021) 4S-DT: self-supervised super sample decomposition for transfer learning with application to COVID-19 detection. IEEE Trans Neural Netw Learn Syst 5(37):952–960
5. Mohammed M, Raghavendra AVK, Lahari GK (2021) Corona virus detection and classification with radiograph images using RNN. In: 5th International conference on trends in electronics and informatics (ICOEI), vol 8(3), pp 1148–1156
6. Sawhney R, Khan LP, Sharma S (2021) An efficient supervised deep learning approach for Covid chest image classification from deep residual networks. In: 4th International conference on recent developments in control, automation & power engineering (RDCAPE), vol 2(10), pp 405–413
7. Varela-Santos S, Melin P (2021) A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks. Inf Sci 545:403–414
8. Nasser N, Emad-ul-Haq Q, Imran M, Ali A, Al-Helali A (2021) A deep learning-based system for detecting COVID-19 patients. In: ICC 2021—IEEE International conference on communications, pp 1–6
9. Mohan N, Kabeer S, Nasir N (2021) Artificial intelligence (AI) in the diagnosis of COVID-19 detection: a review. In: 2021 28th IEEE International conference on electronics, circuits, and systems (ICECS), pp 1–6
10. Devi SK, Amirthavarshini D, Anbukani RS, Harini TKA (2021) Medical decision support system to detect Covid-19 pneumonia using CNN. In: Innovations in power and advanced computing technologies (i-PACT), vol 9(2), pp 6581–6589
11. Hussain BZ, Andleeb I, Ansari MS, Kanwal N (2021) Lightweight deep learning model for automated COVID-19 diagnosis from CXR images. IEEE Int Conf Comput (ICOCO) 2(61):218–223
12. Islam MN, Inan TT, Rafi S, Akter SS, Sarker IH, Islam AKMN (2020) A systematic review on the use of AI and ML for fighting the COVID-19 pandemic. IEEE Trans Artif Intell 1(3):262–270
13. Dong D (2021) The role of imaging in the detection and management of COVID-19: a review. IEEE Rev Biomed Eng 14(1):16–29
14. Visvanathan M, Balasubramanian V, Sathish R, Balasubramaniam S, Sheet D (2021) Assessing lobe-wise burden of COVID-19 infection in computed tomography of lungs using knowledge fusion from multiple datasets. In: 43rd Annual international conference of the IEEE engineering in medicine & biology society (EMBC), vol 6(20), pp 107–115
15. Abdulla HK, Ahmed ZS, Surameery NMS, Rashid RD, Salih SQ (2020) SAARSNet: a deep neural network for COVID-19 cases diagnosis. Int Conf Adv Sci Eng (ICOASE) 13(7):858–866
16. Islam MR, Matin A (2020) Detection of COVID 19 from CT image by the novel LeNet-5 CNN architecture. In: 23rd International conference on computer and information technology (ICCIT), pp 1–5

17. Abdar AK, Sadjadi SM, Zadeh HS, Bashirgonbadi A, Naghibi M (2020) Automatic detection of coronavirus (COVID-19) from chest CT images using VGG16-based deep-learning. In: 27th national and 5th international Iranian conference on biomedical engineering (ICBME), vol 9(1), pp 556–564
18. Jain A, Ratnoo S, Kumar D (2021) Convolutional neural network for Covid-19 detection from X-ray images. Fourth Int Conf Comput Intell Commun Technol (CCICT) 28(4):138–145
19. Dash S, Chakravarty S, Mohanty SN, Pattanaik CR, Jain S (2021) A deep learning method to forecast COVID-19 outbreak. New Gener Comput 39(3–4):515–539
20. Panetta K, Sanghavi F, Agaian S, Madan N (2021) automated detection of COVID-19 cases on radiographs using shape-dependent fibonacci-p patterns. IEEE J Biomed Health Inform 5(19):1985–1992
21. Lucas M, Lerma M, Furst J, Raicu D (2020) Heatmap template generation for COVID-19 biomarker detection in chest X-rays. In: IEEE 20th international conference on bioinformatics and bioengineering (BIBE) vol 1(3), pp 4122–4130
22. Zebari DA, Abdulazeez AM, Zeebaree DQ, Salih MS (2020) A fusion scheme of texture features for COVID-19 detection of CT scan images. In: IEEE International conference on advanced science and engineering (ICOASE), pp 1–6
23. Arora P, Gera S, Kapse VM (2021) Diagnosis of lung segmentation for chest X Ray images using XGBoost. In: IEEE 9th International conference on reliability, infocom technologies and optimization (trends and future directions) (ICRITO), pp 1–6
24. Absar N, Mamur B, Elkhader BA (2022) Development of a computer-aided tool for detection of COVID-19 pneumonia from CXR images using machine learning algorithm. J Radiat Res Appl Sci 19(46):769–776
25. Sanagavarapu S, Sridhar S, Gopal TV (2021) COVID-19 Identification in CLAHE enhanced CT scans with class imbalance using ensembled ResNets. In: IEEE International IOT, electronics and mechatronics conference (IEMTRONICS), pp 1–7
26. Foysal M, Hossain ABM (2021) COVID-19 detection from chest CT images using ensemble deep convolutional neural network. In: 2nd International conference for emerging technology (INCET), vol 3(14), pp 3625–3633
27. Afshar P, Heidarian S, Naderkhani F, Rafiee MJ, Oikonomou A, Plataniotis KN, Mohammadi A (2021) Hybrid deep learning model for diagnosis of Covid-19 Using Ct scans and clinical/demographic data. IEEE Int Conf Image Process (ICIP) 10(4):923–930
28. Nneji GU, Cai J, Jianhua D, Monday HN, Chikwendu IA, Oluwasanmi A, James EC, Mgbejime GT (2021) Enhancing low quality in radiograph datasets using wavelet transform convolutional neural network and generative adversarial network for COVID-19 identification. In: 4th International conference on pattern recognition and artificial intelligence (PRAI) vol 7(12), pp 287–294
29. Mandal S, Khan DA, Jain S (2021) Cloud-based zero trust access control policy: an approach to support work-from-home driven by COVID-19 pandemic. Gener Comput 39(3):599–622
30. Gour M, Jain S (2022) Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network. Biocybernetics Biomed Eng 42(1):27–41

# Design of a Prototypic Mental Health Ontology for Sentiment Analysis of Tweets

**Deepika Chaudhary, Jaiteg Singh, and Puneet Kaur**

**Abstract** Ontologies show great promise in addressing many of the issues of component-based research, and they are a strong example of technology-supported science. A mental health ontology is a collection of concepts used to design a model which can be used by a variety of data storage and retrieval applications. While there have been categories of mental problems for over a century, it is still unclear what terminology like "mental disorder," "disease," and "sickness" mean. Ontologies have been enlisted to help fill in the gaps. Ontologies, for the most part, provide a recognizable language that allows for a computerized bridge between disparate authorities, facilitating translation as research becomes increasingly interdisciplinary. The goal of this research is to figure out what role ontologies have in mental health. The goal of this research is to compile a list of terms connected to mental health and draft a protytypic ontology thereof. The research outcome of this paper is a design of a prototypic ontologies for sentiment analysis of tweets.

**Keywords** Ontology · Mental health · Mental illness · Tweets · Sentiment analysis

## 1 Introduction

Social media is gaining attention and becoming very popular nowadays [1]. Among various platforms, Twitter and Facebook are the most powerful platform which has been used by researchers for opinion mining. Opinions on these platforms spread very fast and have a powerful impact on the mindset of people. Humans usually interact using online platforms which are now enabled by various communication technologies. Since conversation on these platforms disseminate very quickly, therefore, good

---

D. Chaudhary · J. Singh
Chitkara University Institute of Engineering & Technology, Chitkara University, Rajpura, Punjab, India

P. Kaur (✉)
Chandigarh Group of College, Landran, Mohali, India
e-mail: Puneet.4843@cgc.edu.in

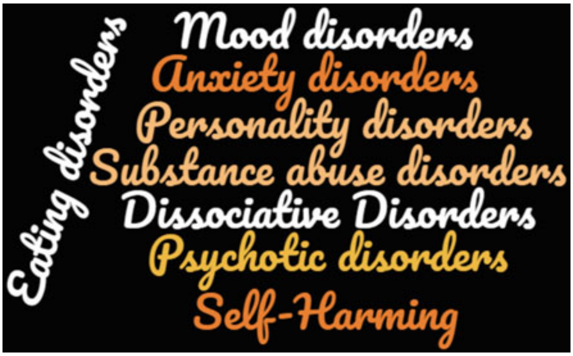
S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_19

and bad information travels at a same speed. However, there is an possibility that valid social media information is of favor as compared to rumors. Twitter can be defined as a micro-blogging site, where the user can post short and long message known as tweets through chats. These tweets can be maximum of 140 characters long. Although most tweets are simple converstion, they are also used to share relevant information [2].

To perform sentiment analysis on the tweets, one requires to extract the words and the concepts which have an powerful impact on the mental health of human being. Ontology is the process of taking the union of the concepts of source and targets ontologies together and expresses the relationships of the concepts in one ontology to the other ontology [3]. In simple words, ontology does not only demonstrate sharable and reused information but can also add new facts in the vocabulary. These well-defined vocabularies are very useful for integrating data coming from various heterogeneous sources. These ontologies act as a collection of entities, and relationships between them are expressed by nodes and edges between these nodes. Ontology is a study of "being" and disturbed among "What is," i.e., the environment life and organization of truth. According to Richards, ontology is the assumptions we formulate regarding the category and natural world and what goes out. Mental health is a part of our mental well-being. Sooner or later, a person always relates their health with mental well-being. The general tendency of many people is to think of mental health as a mental illness. But, it is not true. The concept of mental wellness goes far beyond absence of mental illness. A person's psychological, emotional, and social well-being is referred to as their mental health. There are some common types of illnesses that we face. Figure 1 describes various types of mental health illness.

During pandemics online interactions have become a regular part of our daily lives. Through the messages on social media, human being shares their opinions, thoughts, feelings, on almost every aspect of life. Sometimes, when a human is under pressure or facing any mental illness expresses his views on these sites. These views if inferred properly can protect humans from harming themselves. The objective of the study is to critically review the existing ontologies that can help in data extraction from any social media portal.

**Fig. 1** Types of mental health illness



## 2 Literature Survey

This section focuses on the related work done in the direction of sentiment analysis for different domains. The papers were selected on the basis of year of publication and citation index. Narayanasamy et al. in their paper have mentioned that the ontology-based models work best to extract the emotions. As stated by the authors, ontologies can be constructed to capture the emotions from within the text documents. Through the use of Semantic Web technologies, the data captured thereof can be used for various applications. Mental health [4, 5] is very important from an early stage of our life till the end. The emotional, psychological, and social well-being of a person are all included in their mental health. It has an impact on our feelings, thoughts, and actions. Mental disorder is a global problem and is being faced by various countries. Globally, it is a major threat as one in four or five citizens of any country is diagnosed with a mental disorder that has to be controlled. Therefore, it is of major concern to every country and is a social responsibility as National Advisory Mental Health Council Workgroup, 2010.

After learning through the literature mentioned above, it had been found that mental health ontology can be considered to offer a representation of mental health concepts and interactions. On the basis of which, the data from social media platforms like Twitter can be collected and can be further processed for various research problems related to health care. Therefore, here in this study, we figure out various classes of mental illness and specify a list of keywords. These keywords would act as a metadata for tweets' extraction further analysis.

## 3 Research Methodology for Ontology Creation

To perform the study, we would follow the below-mentioned methodology:

(a) **Understand the Concepts of Mental Health**: In an attempt to solve this problem, it is required that we all should have a deep understanding of terms related to the mental state of a human being. In the table below, we classify mental illness along with the words which indicate the disease (https://edelweisspublications.com/).

(b) **Classifying the terms**: To understand the concepts of mental health, the various categories of mental illness were studied. Table 1 depicts various types of mental illness and the related keywords (Table 2).

(c) **Design of Ontology**

Using these terms, an ontology was designed using protégé as a tool (Figs. 2, 3, 4 and 5).

**Table 1** Literature survey

| Year of publication | Author | Summary |
|---|---|---|
| 2021 | Narayanasamy et al. [4] | The whole spectrum of human emotions connected to COVID-19 subjects may be extracted using the emotion ontology model, which has proven to be accurate and dependable. The suggested system performs satisfactorily when compared to prior lexicon-based emotion algorithms and deep learning techniques |
| 2021 | Ahmed et al. [3] | We analyze users’ sentiment at various time intervals using the COVID-19 tweets’ dataset to aid identify trending topics on Twitter. Sentimental clusters can be deduced from the sensation categories. Using complete sentiment dynamics, we look at a range of experimental findings that show different degrees of diversity in data pertaining to social media communication during the pandemic age |
| 2020 | Barkur et al. [2] | Coronavirus disease, also known as COVID-19, is a new virus outbreak that started in Wuhan, China. The disease now has escalate to practically every country on the planet. The purpose of this paper is to look at how Indians felt after the shutdown |
| 2020 | Rains et al. [5] | The author claims to complete a text analysis in order to provide a preliminary mapping on the early effects of COVID-19 on the mental state of people. Many persons who suffer from mental illnesses are accustomed to being alone. This could lead to greater adaptability and the ability to deal with obstacles |
| 2020 | Gao et al. [6] | During a COVID-19 outbreak in Wuhan, many people were exposed to social media. We look at the widespread presence of mental health issues. China’s government served many health services but pays more attention to depression and anxiety |
| 2020 | Xue et al. [7] | The author explores Twitter messages related to the pandemic diseases COVID-19 and some hashtags like coronavirus, COVID-19, quarantine, etc. We use Latent Dirichlet Allocation (LDA), the machine learning algorithm to identify tweets’ sentiment analysis |

(continued)

**Table 1** (continued)

| Year of publication | Author | Summary |
|---|---|---|
| 2020 | Kejriwal [8] | In this paper, the author analyzes an inventive equipment knowledge graph (KG) that comes under AI technology. KG has a major role in COVID-19 for the doctors, policymakers, epidemiologists, and another domain expert who are trying to dig into the data for deeper analysis of the disaster |
| 2020 | Rabindra Lamsal [9] | During this crisis, everything is affected by this disease COVID-19. These days, people are more interactive in social networking sites like Facebook and Twitter for the fastest news updates. In this paper, the author presents COVID-19. A greater grasp of the geographical and temporal aspects is provided by large-scale geotagged tweets |
| 2019 | Davcheva et al. [10] | The author of this research contributes to the understanding of how mental health users are impacted by healthcare communities. Sentiment analysis on big datasets of user posts was the focus of this work. The objective of this research is to investigate the impact of mood on mental health |
| 2011 | Luxton et al. [11] | The fast growth of smartphones has opportunities for behavioral health care. Smartphones' applications provide useful functionality in clinical health care. It also communicates with real-world two-way communication capabilities |

**Table 2** Categories of mental illness along with keywords

| Type of mental illness/disorder | Meaning | Keywords |
|---|---|---|
| Mood | The person suffering from a mood disorder generally finds it difficult to work inconsistent way. If your emotional state interferes with your routine work and you are either too happy, extremely sad, have trouble with drinking or drugs and suicidal thoughts come to your mind | Merry, joy, mirth, bliss, elation, glee, euphoria, jubilation, rapture, felicity, unhappy, broken-hearted, sorrowful, dejected, regretful, depressed, downcast, miserable, downhearted, grief-stricken, broken heart, saying goodbye, wanting to be left alone, "I'm going to commit suicide," "I wish I had never been born," or "I wish I were dead" |
| Anxiety | Feeling nervous when in normal conditions | Feel like terror, have regular chest pain, palpitations, phobias, fear, angst, isolation, loneliness, sinking |
| Personality | Having an unhealthy pattern of thinking, working, and behaving<br>(Borderline Personality Disorder Test\|BPD Test\|NHS and NZ (zerotoeternity.com) | Intense fear, unstable relationship, sudden anger or aggression, reckless spending |
| Psychotic | When a person finds it difficult to distinguish between real and false, it is a very severe condition | Hallucinations, delusions, off-the-wall behavior, chaotic speech, incoherency, and hearing voices |
| Eating | Illness that occurs due to excessive eating | Irregular eating, excessive eating |
| Self-harming | A person harming himself due to mental distress | Dissociation from the outside world, attention-seeking |
| Substance abuse | The consumption of any substance including alcohol and cocaine, etc | Alcohol, opiates, cocaine, amphetamines, hallucinogens |
| Dissociative | Where a person experiences a disconnection between thoughts, memories, identity, and actions | Dissociation from outside word, personality disorder |

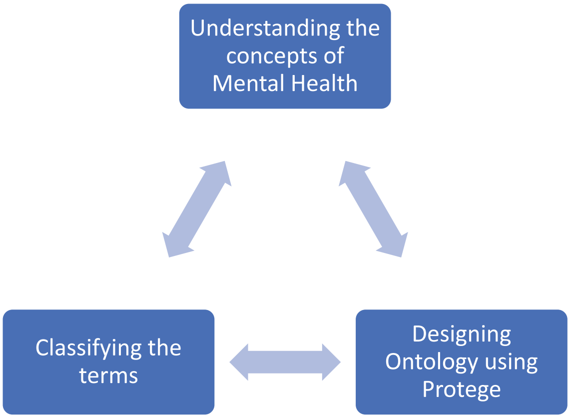


**Fig. 2** Research methodology on mental health

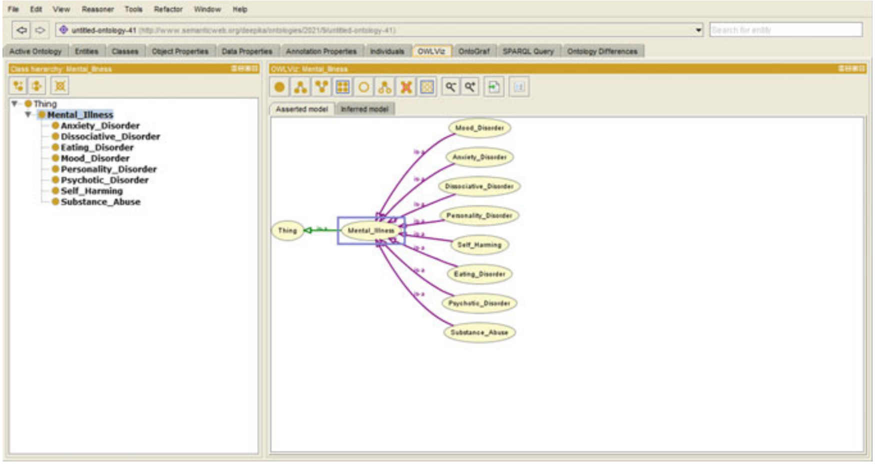


**Fig. 3** Mental illness classes

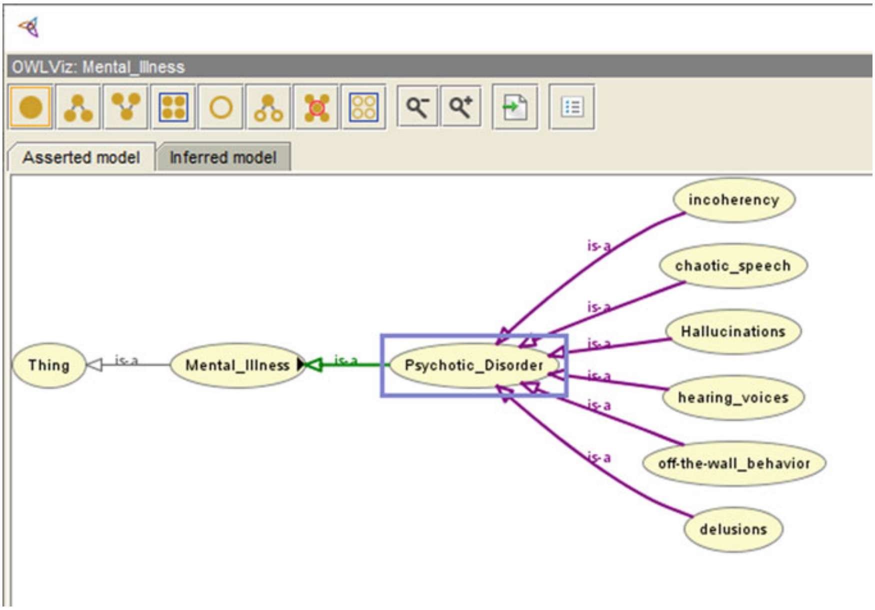


**Fig. 4** Psychotic disorder subclass of mental illness

## 4 Conclusion and Future Scope

In this paper, we have classified various categories of mental illness. We gleaned useful information from the tweets of people who have been diagnosed with depression on Twitter. We recognized the symptoms of depression, self-interest, feelings of irrelevance or guilt, problems with decision-making, suicidal thoughts, mood swings, anxiety, exhaustion, annoyance, and certain signs of fatigue or paranoia for each of the specified groups of depression. A set of keywords or concepts are defined herewith. Using these keywords, a mental illness ontology was created. This ontology would be used as a semantic framework for extracting tweets. We would like to extract more comprehensive information from depressed Twitter users in the future, such as mood swings and changes in behavior with friends and family.

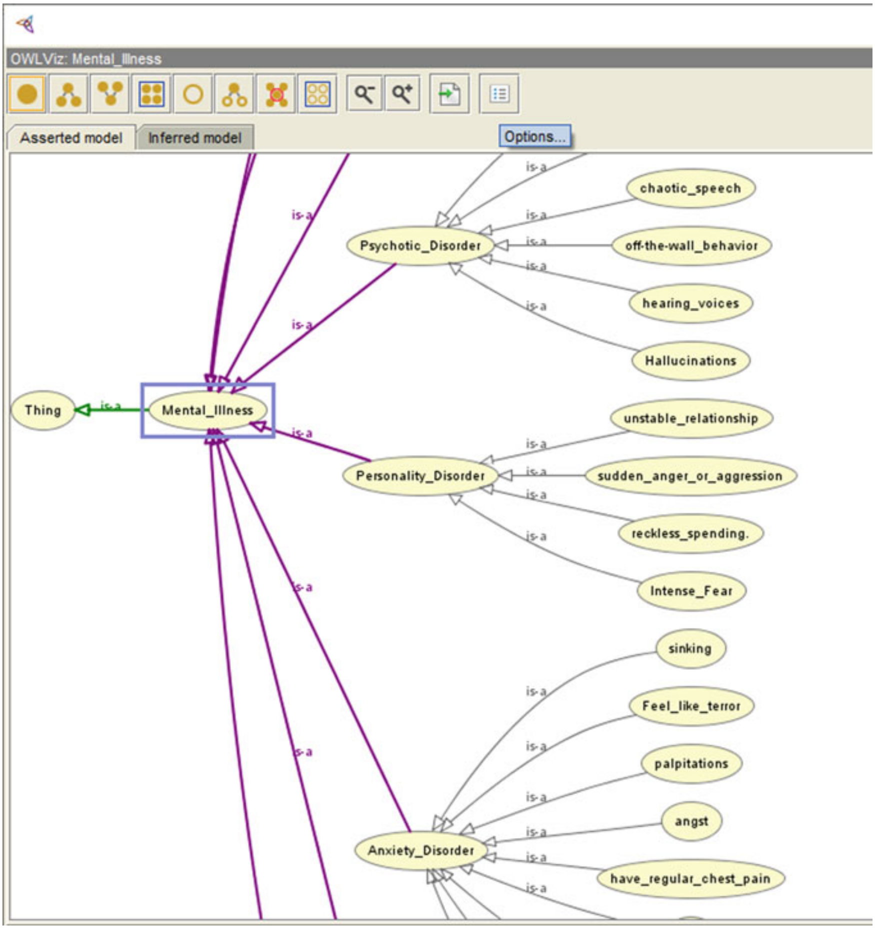


**Fig. 5** Ontology layout

## References

1. Antonakaki D, Fragopoulou P, Ioannidis S (2021) A survey of Twitter research: data model, graph structure, sentiment analysis and attacks. Expert Syst Appl 164:114006
2. Barkur G, Vibha GBK (2020) Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: evidence from India. Asian J Psychiatr 51:102089
3. Ahmed MS, Aurpa TT, Anwar MM (2021) Detecting sentiment dynamics and clusters of Twitter users for trending topics in COVID-19 pandemic. PLoS ONE 16(8):e0253300
4. Narayanasamy SK, Kathiravan Srinivasan SMQ, Chang CY (2021) Ontology-enabled emotional sentiment analysis on COVID-19 pandemic-related Twitter streams. Front. Public Health 9
5. Rains LS, Johnson S, Barnett P, Steare T, Needle JJ, Carr S, Simpson A (2021) Early impacts of the COVID-19 pandemic on mental health care and people with mental health conditions:

framework synthesis of international experiences and responses. Soc Psychiatry Psychiatr Epidemiol 56(1):13–24

6. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, Dai J (2020) Mental health problems and social media exposure during COVID-19 outbreak. PLoS ONE 15(4):e0231924
7. Xue J, Chen J, Hu R, Chen C, Zheng C, Liu X, Zhu T (2020) Twitter discussions and emotions about covid-19 pandemic: a machine learning approach. arXiv preprint arXiv:2005.12830
8. Kejriwal M (2020) Knowledge graphs and COVID-19: opportunities, challenges, and implementation. Special Issue 1-COVID-19
9. Lamsal R (2021) Design and analysis of a large-scale COVID-19 tweets dataset. Appl Intell 51(5):2790–2804
10. Davcheva E, Adam M, Benlian A (2019) User dynamics in mental health forums–a sentiment analysis perspective
11. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM (2011) mHealth for mental health: integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pract 42(6):505

# An Ensemble for Attendance Management with Face Visualization and Recognition Using Local Binary Pattern Histogram

**Guttula Nookaraju, T. Sreeja Reddy, D. Vignesh Kumar, and Madhu Bandari**

**Abstract** Identifying a person and recognizing their individuality has been the most important necessity in any community or organization. So, attendance management is a necessary tool in any environment where maintaining attendance is required. Automation is always an easier way to upgrade the existing approach, in performance, efficiency, and robustness. This project is aimed at developing a less intrusive, cost-effective, and more efficient automated attendance management system. The system is based on face detection and recognition algorithms, which automatically detects a person through a walk-in or an image of an individual or a group. The system comports the given image with the pre-registered data and confronts the lateral procedure. The system uses Local Binary Pattern Histogram Algorithm with Haar cascade classifier to identify the image through its nodal preference to downscale, detect and differentiate the image. The produced data of the attendance is managed and automatically updated in the database and is stored in an excel sheet. Different real-time scenarios were taken into account to evaluate the performance of the system. The project also has measures implemented to handle the threats like spoofing.

**Keywords** Face recognition · Local Binary Pattern Histogram (LBPH) · Biometric systems · Radio Frequency Identification (RFID) · Near Field Communication (NFC)

---

G. Nookaraju · T. Sreeja Reddy (✉) · D. Vignesh Kumar
Vardhaman College of Engineering, Hyderabad, India
e-mail: sreeju.reddy.2000@gmail.com

G. Nookaraju
e-mail: rajuguttula1798@gmail.com

D. Vignesh Kumar
e-mail: vigneshdugyala980@gmail.com

M. Bandari
Department of Data Science and Artificial Intelligence, Faculty of Science and Technology (Icfai Tech), ICFAI Foundation for Higher Education, Hyderabad, India
e-mail: madhu.bandari@ifheindia.org

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_20

# 1 Introduction and Background

Attendance Management systems developed continually over years, owing to specifications involving time, effort, performance, usability, and robustness. The most significant identification processes of the development are the biometric methods, where a person's physiological variants are taken as the basis of classification [1]. It relies on unique specific data about biological traits to work efficiently. So far, biometric methods have proven to be the most effective processes to identify a person, for both identity verification and attendance management. Face recognition is a simplistic way of analysing a persons face on a nodal reference frame to extract the features and recognize a persons identity from the trained and stored data. Though the concept is colloquially specific, the process of identification is varied with the performance and accuracy of the classification algorithm. However, face recognition technology is more flexible when compared to other biometrics in regard with cost, accuracy, portability and is preferred in general owing to its less intrusive procedure. Although face recognition is readily used in many ecosystems, there's a room for improved efficiency in terms of accuracy of the system and efficiency of algorithm. The project focuses on the robustness of this procedure.

## *1.1 Objectives and Scope*

The sole purpose of this project is to build a Face Recognition Biometric Technology that addresses the limitations of manual and other alternative biometric technologies, while providing the perks of its own.

While there are many other existing technologies like RFID and NFC, certain limitations in managing spoofing, flexibility and cost made Face Recognition technology a key to enhancing the improvements.

Although face recognition has limitations of it‘s own regarding accuracy, ideality of identification in different illuminations [2, 3], expressions [4, 5], motion, there is a felicitous chance of improving the efficiency of algorithm and its classification with increased flexibility [5, 6].

# 2 Collation of Various Attendance Management Systems

Originally, attendance was managed manually in registers. It took time, effort and had a maximum possibility of spoofing. Keeping track of a bulky record is inapt to realising information when requested by a legitimate party. So, various other automated technologies were used in place to ease the system.

## 2.1 Near Field Communication

NFC typically uses a tag scanner programme that enables in identifying the uniquely issued tags and relate the information to the user. Being a near field communication system, NFC has a detection range of 4cm or less. Many organizations use NFC installations for their rapidity in updation and accuracy. The project Automated Attendance System using NFC [7] detailed the use of Battery Assisted Passive tags using mobile devices which get battery activated only in the presence of the RF field. However, a typical NFC system doesn't have a mechanism to detect the authenticity of the tag holder, and is not robust against spoofing without additional manual or camera surveillance. Also, in cases of a forgotten tag or mobile, there is no specific backup procedure to address the attendance. NFC technology is submissively well functioning procedure which can still be functionally replaced by biometric methods.

## 2.2 Radio Frequency Identification

RFID is a radio frequency identification procedure that has a larger field length than NFC. A general NFC system works on RFID principle. It uses radio waves to transfer data from RFID tags to identify the host. The project RFID based Student Attendance System [8] proposes a model with a RFID reader programme in a client computer that is always on a vigilance to propagate an electromagnetic field due to a presence of RFID tag in the vicinity. The tag reflects a unique ID to the system to identify the tag owner. However, like NFC, RFID doesn't have a specific feature to authenticate the tag holder. In addition, the installation process is expensive when compared to Face Recognition Technology.

## 2.3 Biometric Systems

ID number or a tag code uniquely identify a person, so do a few physiological traits. Biometric systems rely on a person's physical characteristics to identify them, thus being more specific about the identity. Since this is a science of measuring life, biometry measures uniqueness in the gathered traits of a person and exemplifies the information to the identity. Everyone has a unique fingerprint, palm vein structure, iris pattern, retinal capillary arrangement, thermogram. Each of these physical traits can be used as an identification hinge and require a respective system design. Although all the biometric systems are imperatively efficient, a few subjects provide more accuracy and less chance to knavery.

**Fingerprint Recognition**

Fingerprint recognition is a distinctive process of comparing stored fingerprint patterns to the input, thus relating it to the identity [9]. It is one of the most accurate biometric methods with a low chance of fraud. However, the process requires contact with the fingerprint reader, and hence is a disadvantage in the covering situations of pandemic. Yet, there is a scope for a touchless fingerprint sensor to address this problem, with additional upgrade in installation and cost. Although fingerprint recognition is more accurate, overtime as a person ages, there is chance of fingerprint fading that might lead to a problem involving performance incompatibility in diverse population.

**Retinography**
Retinography identifies a person through a retinal capillary scan, that is distinctive in people. It propagates an infrared beam through the eye which is surrounded by numerous blood filled capillaries that absorb the light more than the muscle tissue and standout to form a pattern. This pattern is unique and helps differentiating people. However, it is non-portable and requires a lot of end user attention involving a little intrusion and direct infrared exposure to the eye, inducing negative opinions among people. Yet when done right, retinal scan is proven to be ideally the most accurate biometric systems.

**Iris Scan**
Like retinography, Iris scan identifies a person through unique pattern in iris by a projection through a visible or near infrared light to obtain a high variant iris picture that helps determining the unique iris pattern. It is a highly secure and accurate biometric technology and needs a specialized hardware. Although studies show that the present recognition system fails while addressing patients with acute iris inflammation known as ocular disease, iris scan is precisely accurate than most other biometrics.

**Face Recognition**
Face Recognition Systems work like any other biometric systems using users face as the value of reference. The face recognition algorithms create a biometric template by downsizing the focus to the users face and detecting, measuring the feature points including their location and structure [10–12]. It typically uses a general webcam for image capture and hence is portable. Although limitations due to illuminations, expressions, facial injuries are persistent, improvement in the classification algorithm can increase the robustness of the system [2, 3, 6]. It is a low cost and flexible technology and can be inherited or incorporated into other biometric methods.

## 3 Proposed Model and Method

The proposed system is a design of a robust, visually digitalized attendance management system that can be reinforced in various levels of corporations.

## *3.1 Model Design*

The model is profusely developed within the system environment with a necessary camera.

**Hardware**: Intel i3-8th Generation or above or relevant processors, 8GB or above RAM, 512MB or above hard disc, HD webcam.
**Software**: Libraries- OpenCV, Tkinter, Pillow, pandas, numpy, datetime, Editor-VS Code, Technology Used- Python 3 or above.

### 3.1.1 Software Development

Face recognition has different working algorithms. Some of them are- Eigenfaces algorithm, Fisher Face algorithm, Open BR, Open Face, Local Binary Patterns Histograms (LBPH). We use LBPH algorithms, since the researches prove it to be effective in resulting true positives than other algorithms [13–16].

**Database Creation**
The first step in the attendance system is the creation of a database of faces that will be used. Different individuals are considered and a camera is used for the detection of faces and the recording of the frontal face. The number of frames to be taken for consideration can be modified for accuracy levels. These image datasets are then saves as yml files and stored in the database along with the registration ID (Fig. 1).

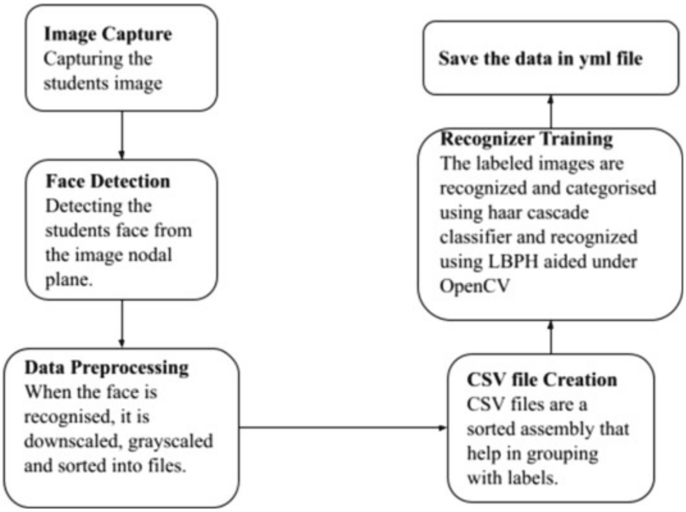


**Image Capture** Capturing the students image

**Face Detection** Detecting the students face from the image nodal plane.

**Data Preprocessing** When the face is recognised, it is downscaled, grayscaled and sorted into files.

**CSV file Creation** CSV files are a sorted assembly that help in grouping with labels.

**Recognizer Training** The labeled images are recognized and categorised using haar cascade classifier and recognized using LBPH aided under OpenCV

**Save the data in yml file**

**Fig. 1** Database creation

### Image Detection and Recognition

LBPH algorithm can recognize human face [17, 18]. It typically makes use of 4 parameters namely Radius(radius around the centre pixel), Neighbours(sample points that build the circular binary pattern, proportional to cost), Grid X(Number of cells in horizontal plane), Grid Y(Number of cells in vertical plane). It typically enrolls the image into recognizer and identifies it in the testing phase.

### Series Working

In general, it starts with the creation of an intermediate image through sliding window concept to give a better representation of the original image with radius and neighbours. The threshold values are compared to the default 8 neighbour values and the neighbour values greater than threshold are set to 1 and others to 0. This composition illustrates the division to further assembly of pixels with gird x and grid y. When comparing the images, the histograms are compared which returns a confidence of measurement that helps recognizing the image [16] (Fig. 2).

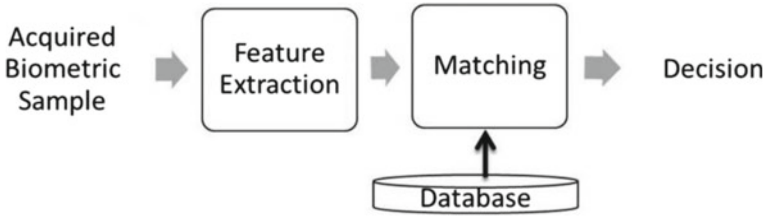
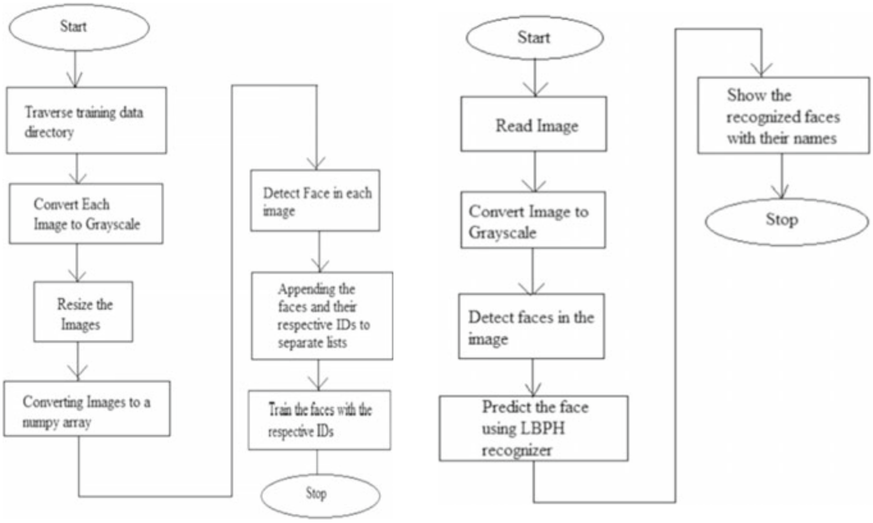


**Fig. 2** Block diagram



(a) (b)

**Fig. 3** **a** Training the recognizer **b** Recognizing the image

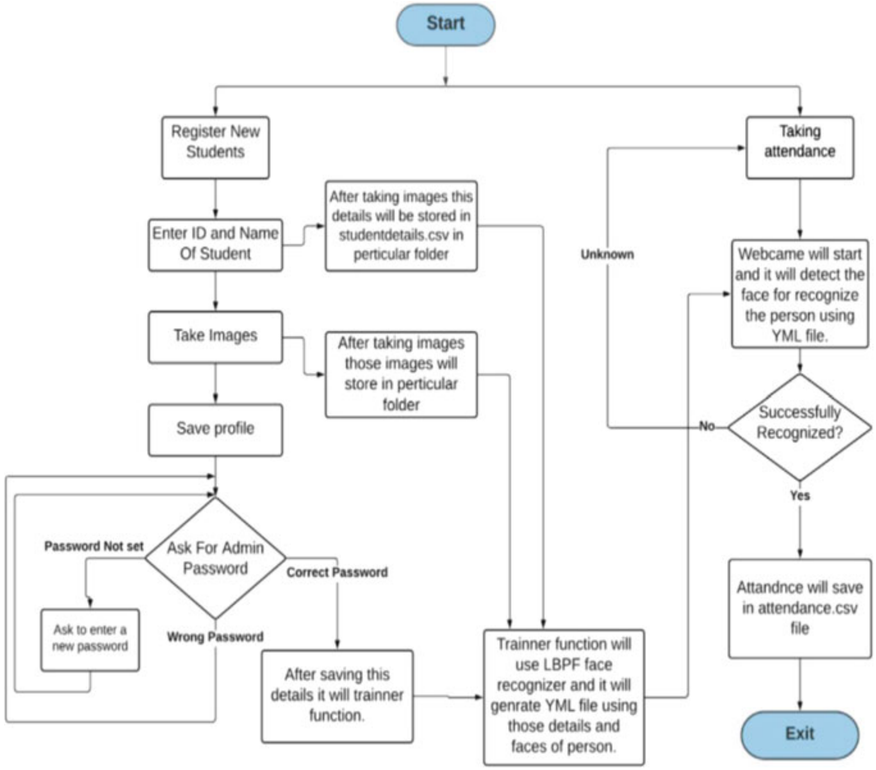


**Fig. 4** Process model

**Training**
Each face undergoes a training in recognizer and is colloquially classified by appending an ID in lists after converting the images into numpy arrays. These datasets are further required during classification (Fig. 3a).

**Recognizing**
During the training, the image reduced histograms act as hash values that uniquely identify the characters based on id’s. These qualities are compared for every new image and this returns a confidence value that helps in identification (Fig. 3b).

**Authentication**
The process of registration and validation is authenticated with a pass-code that can be changed or updated by the management. This verifies the authenticity of the user since the process is supervised by the management. The produced daily attendance is systematically availed to one’s email upon request (Fig. 4).

### *3.2 Methodology*

The attendance is taken in a robust, portable, and user friendly environment with little or no intrusion. When a new user needs to register, images acquired from them are trained and categorized with haar cascades [19] and respective histograms are produced. They are tagged to the ID's of the images in frontal and angular positions and sequenced into the yml files. When the user's picture is uploaded or captured, the lpb recognizer compares the deduced histogram to the database of stored yml files and draws a confidence value to help recognize the person [8]. LBPH proves flexible to various positions and illuminations, though not ideally accurate. The recognizer can recognize multiple faces in a snapshot and tag them with their id‘s [20, 21]. The attendance is simultaneously stored in an excel sheet that can only be available to the management and requires their permission for access. It is simple linear process without much manual intrusion. The process is wildy specific to the arena of purpose and installation and can be apt to being portable in a handy camera or a bluetooth device.

## 4 Results

The following snapshots include the process of registering a new user and identifying him during attendance. The first image shows the application view with no registered users. Now as the process proceeds, the user requires to get the register access through a secured channel using the password under the admin supervision. The recognizer then classifies the image and the user gets recognized. The last image gives the resultant excel sheet of the session‘s attendance (Figs. 5, 6, 7, 8, 9, 10, 11 and 12).

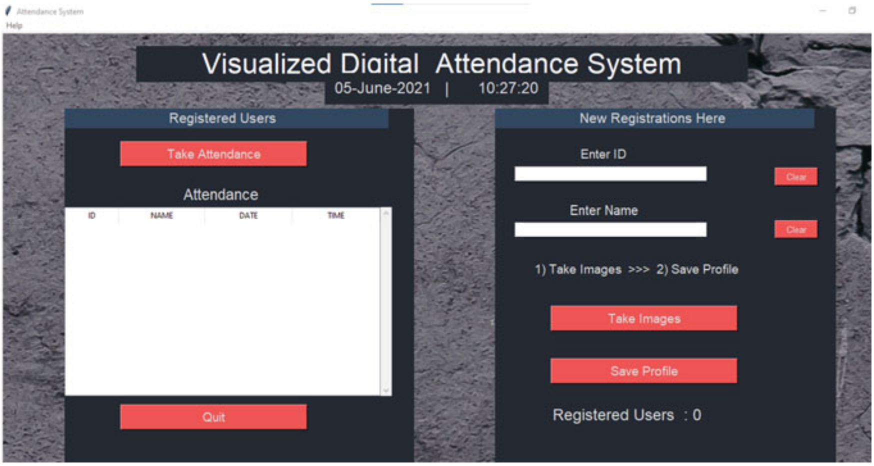


Fig. 5 Application view

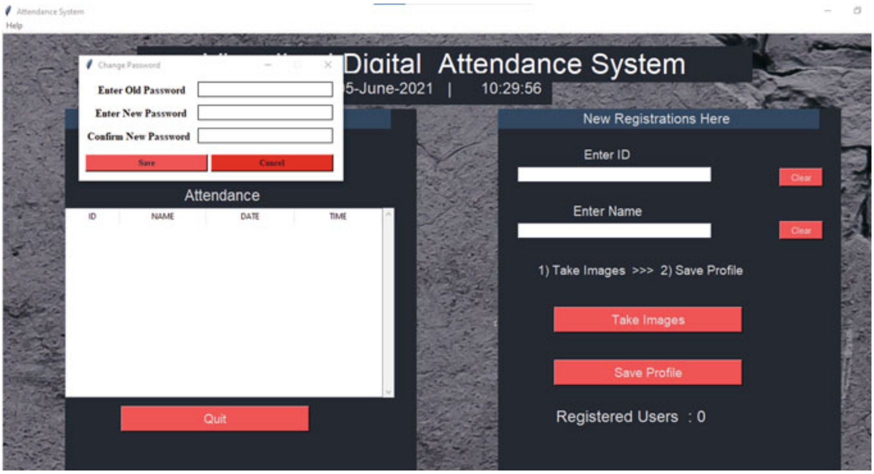
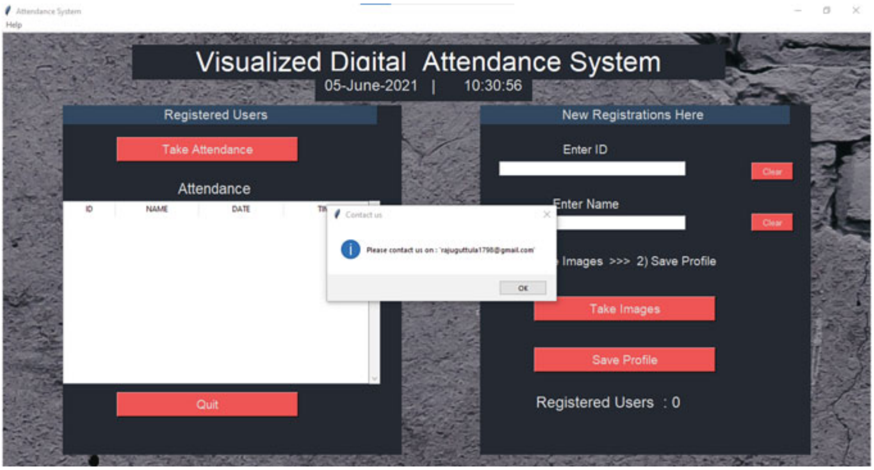
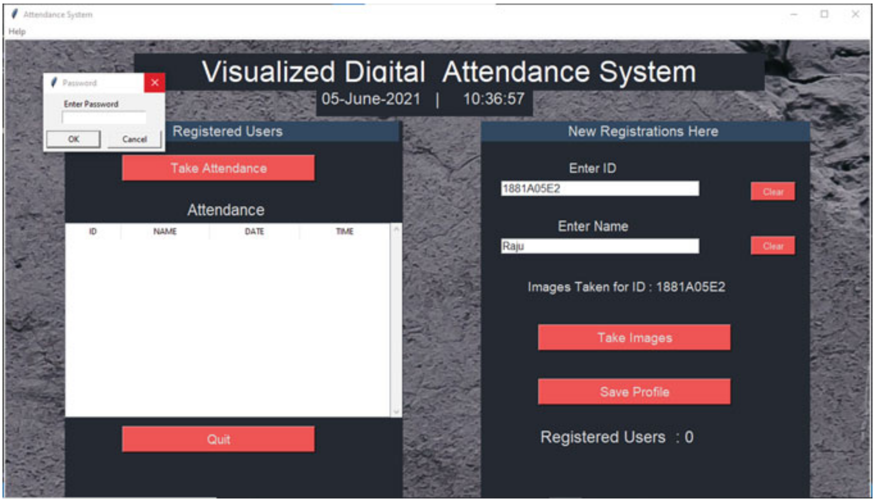


Fig. 6 Authentication



Please contact us on : 'rajuguttula1798@gmail.com'

Fig. 7 Redirect to management contact information

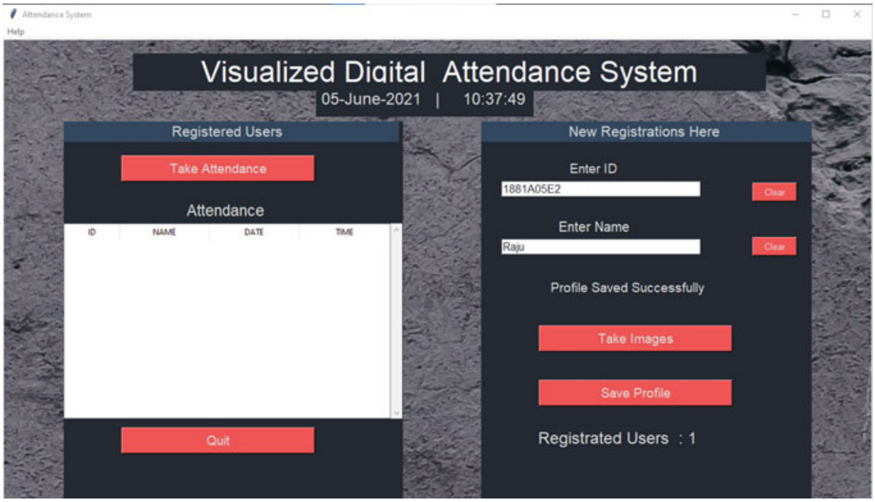


Enter ID: 1881A05E2

Enter Name: Raju

Images Taken for ID : 1881A05E2

Fig. 8 Registering with pass-code



Visualized Digital Attendance System

05-June-2021 | 10:37:49

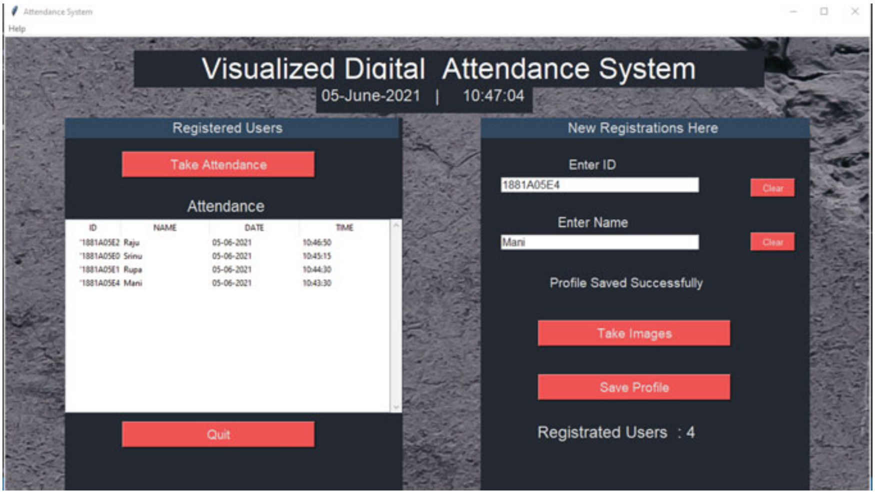
Enter ID: 1881A05E2

Enter Name: Raju

Profile Saved Successfully

Registrated Users : 1

**Fig. 9** View after registration



Visualized Digital Attendance System

05-June-2021 | 10:47:04

| ID | NAME | DATE | TIME |
|---|---|---|---|
| '1881A05E2 | Raju | 05-06-2021 | 10:46:50 |
| '1881A05E0 | Srinu | 05-06-2021 | 10:45:15 |
| '1881A05E1 | Rupa | 05-06-2021 | 10:44:30 |
| '1881A05E4 | Mani | 05-06-2021 | 10:43:30 |

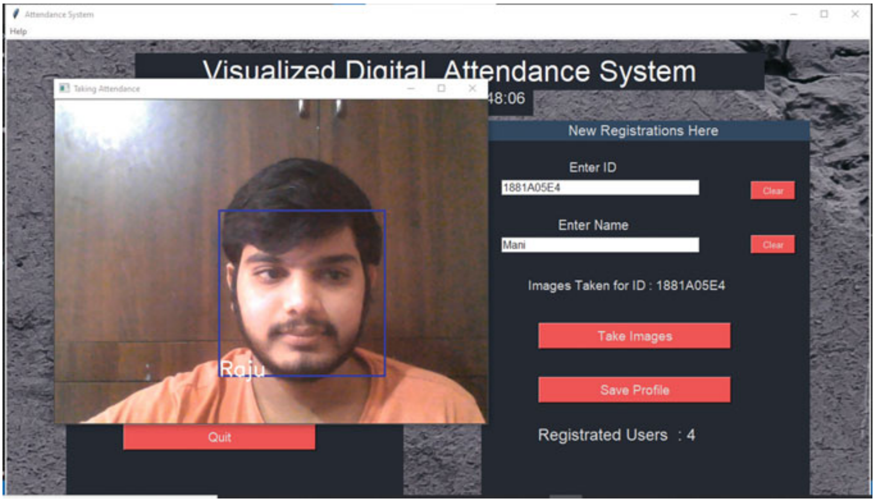
Enter ID: 1881A05E4

Enter Name: Mani

Profile Saved Successfully

Registrated Users : 4

**Fig. 10** Image capture for training



Enter ID: 1881A05E4

Enter Name: Mani

Images Taken for ID : 1881A05E4

Registrated Users : 4

**Fig. 11** User face recognition

Enter ID: 1A05E4

Images Taken for ID : 1881A05E4

Registrated Users : 4

**Fig. 12** Attendance excel sheet

## 5 Concerns and Scope

Although face recognition system is flexible than other systems in terms of accuracy relevance with cost, portability, ease in use, certain conditions involving detection of face in different illuminations [3], under masks, motion and recognizing identical faces(the twin problem) [22] do cause limitations to the system at different levels [23]. However, enhancements in algorithms and cost management did show improvement under various levels considering the use of high quality infrared cameras for detection and psychological variances that effect expressions in twin problems, there is a scope for advancement. Apart from that, this system can be incorporated into various streams of organization like medical, government, military and also virtual conferences. Face recognition is all about accuracy in different conditions and hence there‘s always a plenty of room for improvement.

**Acknowledgements** We solemnly acknowledge Dr.Sachi Nandan Mohanty, Associate Professor at Vardhaman College of Engineering for a constant guidance and support throughout the research process and for helping us through a proper analysis of various interrelated studies.

## References

1. Genovese A, Piuri V, Scotti F (2014) Biometric systems, part of the advances in information security book series, vol 60. ADIS
2. Ravi S, Rethinakumari M (2011) Improving the efficiency of face recognition system for images with different illumination conditions
3. Zeenathunisa S (2012) Next generation face recognition system under various illuminations
4. Lowe DG (2004) Distinctive image features from scale-invariant keypoints
5. Zeng Z, Pantic M, Roisman GI, Huang TS (2009) A survey of affect recognition methods: audio, visual, and spontaneous expressions. IEEE Trans Pattern Anal Mach Intell 31(1):39–58. https://doi.org/10.1109/TPAMI.2008.52
6. Paul KC, Aslan S (2021) An improved real-time face recognition system at low resolution based on local binary pattern histogram algorithm and CLAHE. Opt Photonics J 11(4)
7. Rokade Y, Shetty P, Wankhade S (2017) Automated attendance system using NFC
8. Hussain E, Dugar P, Deka V, Hannan A (2014) RFID based student attendance system
9. Sarfraz M (2021) Introductory chapter: on fingerprint recognition
10. Joseph SB, Dada EG, Misra S, Ajoka S (2022) Parallel faces recognition attendance system with anti-spoofing using convolutional neural network
11. Suresh V, Dumpa SC, Vankayala CD, Aduri H, Rapa J (2019) Facial recognition attendance system using python and OpenCv
12. Shah K, Bhandare D, Bhirud S (2020) Face recognition-based automated attendance system. In: Advances in intelligent systems and computing book series (AISC, vol 1165)
13. Jagtap AM, Kangale V, Unune K, Gosavi P (2019) A study of LBPH, eigenface, fisherface and haar-like features for face recognition using OpenCV. In: 2019 International conference on intelligent sustainable systems (ICISS), pp 219–224. https://doi.org/10.1109/ISS1.2019.8907965

14. Kadir K, Kamaruddin MK, Nasir H, Safie SI, Bakti ZAK (2014) A comparative study between LBP and haar-like features for face detection using OpenCV. In: 2014 4th International conference on engineering technology and technopreneuship (ICE2T), pp 335–339. https://doi.org/10.1109/ICE2T.2014.7006273
15. Kaur P, Mittal P (2021) A comparative study of LBPH, SIFT and SURF algorithms for face recognition task
16. Bhavsar N, Shah S (2020) Face detection algorithms: a comparison
17. Abuzneid MA, Mahmood A (2018) Enhanced human face recognition using LBPH descriptor, multi-KNN, and back-propagation neural network. IEEE Access 6:20641–20651. https://doi.org/10.1109/ACCESS.2018.2825310
18. Wang L, Siddique AA (2020) Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology. https://doi.org/10.1177/0020294020932344
19. Padilla R, Costa Filho CFF, Costa M (2012) Evaluation of haar cascade classifiers for face detection
20. Muyambo P (2018) An investigation on the use of LBPH algorithm for face recognition to find missing people in Zimbabwe
21. Khairnar V, Khairnar CM (2021) Face recognition based attendance system using Cv2
22. Patil K, Bojewar S (2015) A survey on face recognition of identical twins
23. Abayomi-Alli A, Atinuke O, Onashoga SA, Misra S, Arogundade O, Abayomi-Alli O (2020) Facial image quality assessment using an ensemble of pre-trained deep learning models (EFQnet). In: 2020 20th international conference on computational science and its applications (ICCSA), pp 1–8. https://doi.org/10.1109/ICCSA50381.2020.00013
24. Andrejevic M, Sewyn N (2019) Facial recognition technology in schools: critical questions and concerns
25. Pradeep Kumar GH, Ashwini M, Divya GN, Manjushree BN (2015) Multiple face detection and recognition in real-time using open CV. In: Int J Eng Res Technol (IJERT) NCRTS 3(27)

# Healthcare Question–Answering System: Trends and Perspectives

**Shrutikirti Singh and Seba Susan**

**Abstract** Question–answering (QA) system has been a major breakthrough in the field of medical science. An abundant amount of research work has been done in the field of automated text-based QA agents also called chatbots using different neural networks. There has been a steady rise in the number of QA agents in the medical domain in recent times since the Covid-19 pandemic started in December 2019. People have resorted to online consultation to resolve their health issues due to lockdowns and movement restrictions. In this paper, we discuss contemporary healthcare chatbots that answer user queries related to various health issues. There are two types of such agents based on the response generation—first, the conversational type empathetic chatbots employed for psychological or mental health counseling, and second, the medical QA chatbots having domain-specific medical knowledge. In this paper, we concentrate on the second type of chatbots that are used for healthcare question–answering. We outline the limitations of general chatbots and discuss the shifting trend toward the development of more personalized healthcare chatbots in the near future.

**Keywords** Healthcare chatbots · Question–answering agent · Deep learning · Medical domain knowledge

## 1 Introduction

Question–answering (QA) system is an attractive research area for many researchers working in the fields of information retrieval and natural language processing (NLP). QA agents are virtual conversational agents with domain-specific knowledge [1]. Users generally prefer QA agents to document reading since sometimes users want precise answers to questions in a very short time. A QA system, also known as chatbot, basically consists of three main constituents [2]: (i) query processing whose crucial part is the question classification, (ii) document processing whose crucial

---

S. Singh · S. Susan (✉)
Department of Information Technology, Delhi Technological University, Delhi, India
e-mail: seba_406@yahoo.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_21

part is the information retrieval, and (iii) answer processing whose crucial part is the answer extraction. A simple QA agent can be implemented using natural language processing techniques such as computing the term frequency [3] and using it to predict the target word in response to the input query based on historical data [4]. Such chatbots generate stereotypical, though fast, responses [5].

The most popular model used in the last decade to construct a QA system uses the long short-term memory (LSTM) [6]. This model is called a sequence-to-sequence (seq2seq) model proposed in 2014 by Cho et al. [7] and Sutskever et al. [8]. Convolutional neural networks are used sometimes in place of LSTM to create faster models [9]. A LSTM encoder and decoder are required to model the input query and to generate the output response, respectively [10]. Improvisations in this basic model include the introduction of attention mechanisms [11–13] and the substitution of LSTMs with transformers that were introduced by Vaswani et al. [14]. In this survey paper, we review the latest QA agents or chatbots in health care. Healthcare chatbots are of two types: one that can hold empathetic conversations with the user for mental health issues and the other for answering specific health-related queries that require domain-specific medical knowledge; the latter is the topic of study in this paper. We study the challenges in the deployment of well-known seq2seq models for response generation in the medical domain. We also outline the limitations of general chatbots and discuss the shifting trend toward the development of more personalized healthcare chatbots in the near future. The hierarchical structure of study in this paper is given in Fig. 1.

The main contributions of this paper are as follows, along with the sections in which they appear:

1. Detailed discussion on state-of-the-art text-based medical question–answering system (Sect. 2).

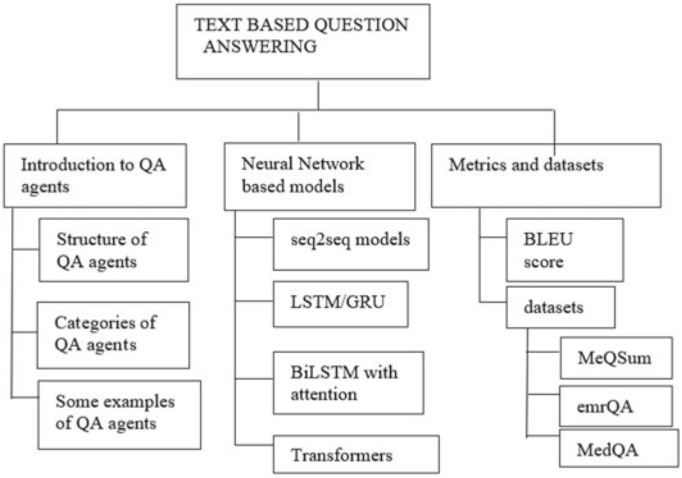


**Fig. 1** Hierarchical structure of this study

2. Categorizing different neural networks applied to various QA tasks (Sect. 3).
3. Conclusion and future scope: highlights challenges for further study in QA systems (Sect. 4).

## 2 A Review of QA Agents Used in Health Care

Question–answering is a well-explored task in the field of natural language processing. It was reintroduced in the early 2000s with the Text Retrieval Conference (TREC) Question–Answering track [15], and other evaluation-based competitions followed suit. There are several applications of text-based question–answering system in the medical, entertainment, and social media domains. This section reviews some traditional state-of-the-art QA agents used in health care.

### 2.1 Types of QA Agents

There are basically two major types of medical QA systems: the first type is fact or knowledge-based that makes use of structured data sources and the second type is based on information retrieval that makes use of search engine and then NLP with machine learning to generate appropriate responses [16]. Factoid question–answering system is a category of QA agents in which the answer is a specific fact, such as a location or a person [17]. These models locate the answer by determining the type of answer during a question categorization step and then collecting a set of possible documents or passages via a search engine. Using normal information retrieval techniques, a response is extracted from a set of candidate responses. In most cases, strategies like these have been used to extract answers by pattern matching [18], question similarity matching using simple word-based features [19], or simple logic-based systems [20]. Web-based QA agents rely on the redundant data available on the web to rewrite the query for efficient answer mining [21, 22]. The development of multi-lingual chatbots is the need of the hour, considering that people across cultures and nations are connected through the internet. Such systems often require a machine translation unit in addition to the QA model. Some multi-lingual healthcare chatbots have been developed for the Hindi [23] and Arabic [24] languages. Multi-lingual biomedical thesaurus such as UMLS is proved to be helpful in such cases [25]. Our survey paper concentrates more on the medical QA agents that are domain-specific, and some of these chatbots became highly popular during the Covid-19 pandemic which has raised health concerns all over the world.

## *2.2 Some Recent Healthcare QA Agents*

In this sub-section, we describe some notable, recently introduced QA agents in the medical domain. There are two types of healthcare chatbots based on the response generation—the empathetic chatbots employed for psychological counseling [26, 27], and the medical QA chatbots with domain-specific medical knowledge [28–30]. One example of a recent healthcare chatbot is Florence [31] that provides diagnosis based on user's symptoms. It makes use of a dataset of illnesses and their symptoms and is designed using the RASA framework. MedQA is another example of a medical question–answering system which aims to answer questions in clinical medicine using the knowledge extracted from large-scale documents [32]. MedQA uses a CNN-based neural network model. Smart mental health care involves combining inputs from social media data, internet-of-things, medical knowledge available on the web, and the electronic health records of patients, as per the survey in [33]. We next discuss the distinctive architectures of some recently introduced healthcare chatbots.

(a) Text messaging-based medical diagnosis using NLP and fuzzy logic [34]: a novel application of a QA agent for online diagnosis. The authors used a fuzzy support vector machine (SVM) for diagnosis. They developed a hybrid neural network by combining LSTM and CNN for sentence embedding. To evaluate the performance, bilingual evaluation understudy (BLEU) score was used. The block diagram of this model is shown in Fig. 2.

(b) HHH: a medical chatbot based on knowledge graphs and text similarity model [35]: the medical data are obtained from the internet. A hierarchical Bi-LSTM attention model is used to look out for the most similar question from a large QA dataset. A semantic similarity based on the Manhattan distance is used to compute the similarities between the user query and the questions in the database. The response of the most similar question is output as the response of the chatbot to the user query. Training and testing are done on a medical subset of the Quora duplicate questions dataset. The process flow of HHH is shown in the block diagram in Fig. 3.

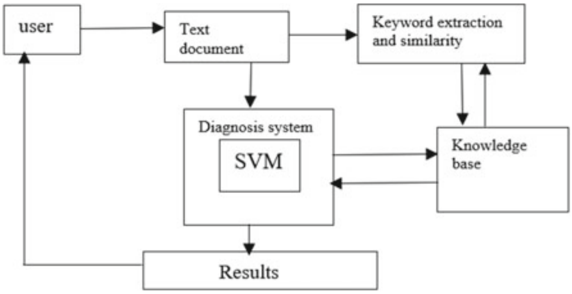


**Fig. 2** Block diagram of the QA model using NLP and fuzzy SVM in [34]

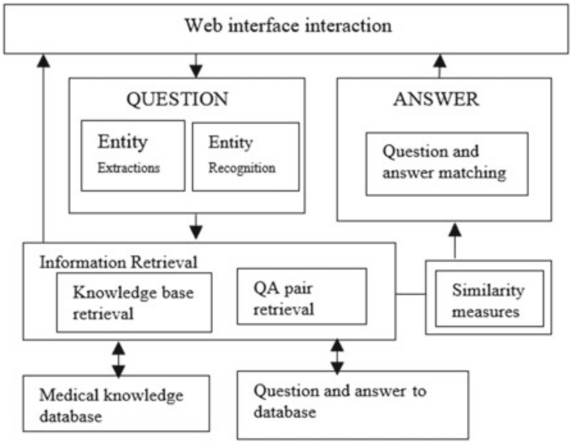


**Fig. 3** Block diagram of the HHH QA model in [35]

(c) CARO [36]: an empathetic conversational chatbot that also provides medical advice for people with depression. Harilal et al. developed a chatbot application in 2020 which holds dual functionality, i.e., medical advice generation and empathetic dialogue generation; the choice between the two depends on the user query and is decided with the help of a binary classifier in the initial stage. The first model is trained on data which contains medical questions and answers, and the second model is trained on data with empathetic dialogue responses. All classifiers have an LSTM-based architecture and use teacher forcing principle for training. The process flow for CARO is shown in Fig. 4.

(d) AskHERMES [37]: a medical question–answering agent that answers complex clinical questions. The data which were collected are preprocessed to calculate semantic similarity. In the third module, i.e., question analysis, the questions are classified under different categories for making information retrieval easy. The final module deals with text-based summarization and answer representation. The process flow is shown in Fig. 5.

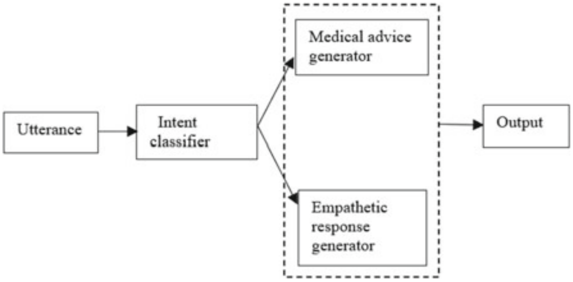


**Fig. 4** Block diagram of the CARO QA model in [36]

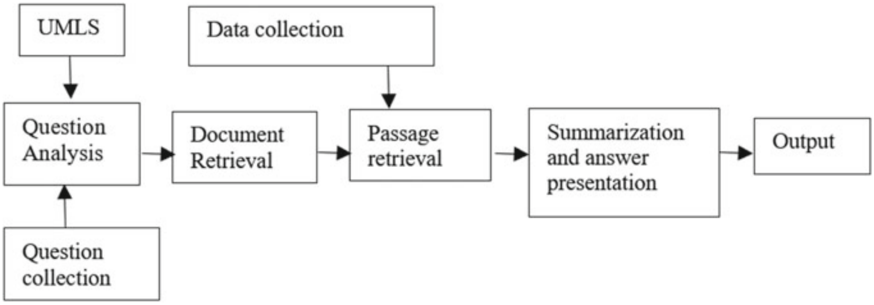


**Fig. 5** Block diagram of the AskHERMES QA model in [37]

## 3 Neural Network-Based Healthcare QA Systems—A Discussion on the State of the Art

In this section, we compare the architectural and functional differences of some state-of-the-art neural network-based question–answering (QA) agents used in the medical domain. By using neural networks, it is possible to process large number of documents with great accuracy and generate responses that are not stereotypical or repetitive [12, 27]. The problem is interpreted as a sequence-to-sequence mapping which is achieved, popularly, using LSTM encoder–decoder modules or seq2seq models [7]. Other than QA systems/chatbots, seq2seq models are also popularly used for neural machine translation and text summarization.

Our contribution in this work is the compilation of the most recent works on QA agents based on state-of-the-art neural networks, i.e., CNN, LSTM, Bi-LSTM, Bi-LSTM with attention and transformers. Bi-LSTM with attention mechanism has been proved to be the most popular neural network in the last decade for constructing QA agents [12, 38]. The attention mechanism is incorporated between the encoder and decoder modules. The introduction of transformers with multi-head self-attention in 2017 [14] has revolutionized the field of natural language understanding and natural language generation. Transformers have started replacing Bi-LSTM with attention in several state-of-the-art seq2seq models including QA systems. Some state-of-the-art neural network-based medical QA agents are summarized in Table 1.

As observed from Table 1, the latest state of the art in sequence modeling is the transformer with multi-head self-attention introduced by Vaswani et al. [14]. Transformers are now replacing LSTMs in all sequential modeling tasks. The multi-head self-attention mechanism of the transformer allows it to analyze different parts of the sentence at the same time, enabling it to capture context at several scales. This is an improvement over the Bi-LSTM with self-attention that incorporates a single attention layer between the encoder and decoder modules.

**Table 1** Some state-of-the-art neural network-based QA agents for health care

| Author, year | Method | Approach | Dataset | Contribution |
|---|---|---|---|---|
| Cao et al. [37] | AskHERMES | Question classification using SVM and keyword identification, followed by text summarization and answer presentation | Medline abstracts, PubMed, eMedicine, clinical guidelines, Wikipedia | The answers are presented in the form of clusters or ranked lists or a combination of both. Clustered answers were found apt for questions covering several topics |
| Ni et al. [39] | CliniQA | Question analyzer followed by retrieval of candidate articles followed by answer generator | Trip answers' website containing 6382 questions | Integration of probabilistic candidate scores to identify the evidences that qualify as the answer |
| Oh et al. [40] | Psychiatric counseling for mental health care | Gated recurrent unit (GRU)-based similarity model | Korean Wikipedia, Namuwiki, and news articles | Chatbot provides conversational service to diagnose psychiatric disorders through the combination of spatiotemporal context analysis and emotion recognition |
| Zhang et al. [32] | MedQA | SeaReader, a LSTM-based network with dual-path attention | Medical reports and examination reports of patients, medical textbooks and articles | Integration of information from multiple text documents that are retrieved from a large-scale database as being most semantically relevant to a question–answer pair |
| Zhou et al. [41] | DUT-BIM | Transformer-based seq2seq model | BioNLP 2019 QA task | A BioBERT transformer model is proposed to extract semantic relation between the question and the answer |

(continued)

**Table 1** (continued)

| Author, year | Method | Approach | Dataset | Contribution |
|---|---|---|---|---|
| Bao et al. [35] | HHH: chatbot based on knowledge graphs and text similarity | Attention mechanism in LSTM-based models is used to compute a semantic similarity score for retrieving the answer | Medical subset of the Quora duplicate questions dataset | Using different models like hierarchical bidirectional LSTM with attention, Siamese LSTM, BERT to calculate the semantic similarity between user query and questions in the database |
| Harilal et al. [36] | CARO | LSTM-based models trained separately on two different datasets | Empathetic dialogues' datasets and medical question–answer dataset | This model provides dual functionality, i.e., medical advice and empathetic response facilitated through a binary intent classifier in the initial stage |
| Ren et al. [42] | CGMVQA: medical visual question–answering | Transformer-based seq2seq model | ImageCLEF 2019 VQA-Med dataset | Three types of text embeddings used for the transformer model, along with ResNet152 to process the input medical images |
| Soni and Roberts [43] | Pre-training and fine-tuning of transformers | Transformer-based seq2seq model | CliCR and emrQA | Initial fine-tuning improves the performance of the QA system |
| Alzubi et al. [44] | COBERT: Covid-19 QA system using BERT | TF-IDF for scoring documents and DistilBERT for scoring question–answer pairs | Covid-19-related literature (from CORD-19 dataset) to find top match with user query | Document retriever using term frequency count followed by finding suitable answers in the retrieved documents using DistilBERT transformer |
| Yadav et al. [45] | Question-aware transformer model | Question-based abstractive text summarization using transformer | MeQSum dataset | Presents a question-driven text summarization of the clinical text as the answer to the user query |

## 4 Conclusions and Future Scope

This study is instituted with the aim of providing an overview of healthcare QA systems and discusses the base models of QA agents used in medical domain. Moreover, we discussed the existing text-based QA models along with the corresponding neural networks they adopted. The role of the attention mechanism is emphasized in our study with both LSTM- and transformer-based models incorporating some form of attention mechanism in their architectures. The attention mechanism captures the context in the input query which helps in generating a relevant response. We conclude from our study that deep learning approaches like Bi-LSTM with attention mechanisms and the transformer models give faster results and more human-like responses as compared to other QA systems. The development of multi-lingual chatbots is the need of the hour, considering that people across cultures and nations are connected through the internet. We focus on general purpose healthcare chatbots in this study. The limitations of general healthcare chatbots over personalized healthcare chatbots like kBot [46] are that they do not cater to specific requirements of a particular user. Personalized chatbots track the patient's health over time and learn interactively through a sequence of question–answers over time. Another future direction of research is the development of specialized QA agents that can help in the diagnosis of the disease. Personalized healthcare QA systems that provide diagnosis and treatment of Covid-19 and post-Covid conditions can be envisaged in the near future. Development of large-scale QA datasets, such as the recently introduced RxWhyQA [47] for clinical queries having multiple answers, should be initiated for the same.

## References

1. Gao J, Galley M, Li L (2019) Neural approaches to conversational AI: question answering, task-oriented dialogues and social chatbots. Now Foundations and Trends
2. Allam AMN, Haggag MH (201) The question answering systems: a survey. Int J Res Rev Inf Sci (IJRRIS) 2(3)
3. Susan S, Keshari J (2019) Finding significant keywords for document databases by two-phase maximum entropy partitioning. Pattern Recogn Lett 125:195–205
4. Cerezo J, Kubelka J, Robbes R, Bergel A (2019) Building an expert recommender chatbot. In: 2019 IEEE/ACM 1st international workshop on bots in software engineering (BotSE). IEEE, pp 59–63
5. Ciechanowski L, Przegalinska A, Magnuski M, Gloor P (2019) In the shades of the uncanny valley: an experimental study of human–chatbot interaction. Futur Gener Comput Syst 92:539–548
6. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
7. Cho K, van Merriënboer B, Gulçehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1724–1734
8. Sutskever I, Vinyals O, Le QV (2014) Sequence to sequence learning with neural networks. Adv Neural Inf Process Syst 27

9. Mallick R, Susan S, Agrawal V, Garg R, Rawal P (2021) Context-and sequence-aware convolutional recurrent encoder for neural machine translation. In: Proceedings of the 36th annual ACM symposium on applied computing, pp 853–856
10. Lokman AS, Ameedeen MA (2019) Modern chatbot systems: a technical review. In: Proceedings of the future technologies conference. Springer, Cham, pp 1012–1023
11. Bahdanau D, Cho KH, Bengio Y (2015) Neural machine translation by jointly learning to align and translate. In: 3rd International conference on learning representations, ICLR 2015
12. Goe, R, Vashisht S, Dhanda A, Susan S (2021) An empathetic conversational agent with attentional mechanism. In: 2021 International conference on computer communication and informatics (ICCCI). IEEE, pp 1–4
13. Luong M-T, Pham H, Manning CD (2015) Effective approaches to attention-based neural machine translation. In: Proceedings of the 2015 conference on empirical methods in natural language processing, pp 1412–1421
14. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. Adv Neural Inf Process Syst 30
15. Clifton T, Teahan WJ (2004) Bangor at TREC 2004: question answering track. In: TREC
16. Cairns BL, Nielsen RD, Masanz JJ, Martin JH, Palmer MS, Ward WH, Savova GK (2011) The MiPACQ clinical question answering system. In: AMIA annual symposium proceedings, vol 2011. American Medical Informatics Association, p 171
17. Iyyer M, Boyd-Graber J, Claudino L, Socher R, Daumé H III (2014) A neural network for factoid question answering over paragraphs. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 633–644
18. Soubbotin MM, Soubbotin SM (2002) Use of patterns for detection of likely answer strings: a systematic approach. In: TREC, vol 52, p 90
19. Mihalcea R, Moldovan D (2000) Semantic indexing using WordNet senses. In: ACL-2000 workshop on recent advances in natural language processing and information retrieval, pp 35–45
20. Furbach U, Glöckner I, Helbig H, Pelzer B (2010) Logic-based question answering. KI-Künstliche Intelligenz 24(1):51–55
21. Dumais S, Banko M, Brill E, Lin J, Ng A (2002) Web question answering: is more always better? In: Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval, pp 291–298
22. Jain S, Jain V, Balas VE (eds) (2021) Web semantics: cutting edge and future directions in healthcare. Academic Press
23. Dhanda A, Goel R, Vashisht S, Susan S (2021) Hindi conversational agents for mental health assistance. Int J Appl Res Inf Technol Comput (IJARITAC) 12(1–3):12–20
24. Shawar BA, Atwell E (2009) Arabic question-answering via instance based learning from an FAQ corpus. In: Proceedings of the CL 2009 international conference on corpus linguistics, vol 386, no 1. UCREL, pp 1–12
25. Athenikos SJ, Han H (2010) Biomedical question answering: a survey. Comput Methods Programs Biomed 99(1):1–24
26. Rashkin H, Smith EM, Li M, Boureau Y-L (2019) Towards empathetic open-domain conversation models: a new benchmark and dataset. In: Proceedings of the 57th annual meeting of the association for computational linguistics, pp 5370–5381
27. Goel R, Susan S, Vashisht S, Dhanda A (2021) Emotion-Aware transformer encoder for empathetic dialogue generation. In: 2021 9th International conference on affective computing and intelligent interaction workshops and demos (ACIIW). IEEE, pp 1–6
28. Lally A, Bagchi S, Barborak MA, Buchanan DW, Chu-Carroll J, Ferrucci DA, Glass MR et al (2017) WatsonPaths: scenario-based question answering and inference over unstructured information. AI Mag 38(2):59–76
29. He X, Zhang Y, Mou L, Xing E, Xie P (2020) Pathvqa: 30000+ questions for medical visual question answering. arXiv preprint arXiv:2003.10286
30. Guo Q-L (2010) A novel approach for question answering and automatic diagnosis based on pervasive agent ontology in medicine. Int J Intell Syst 25(7):655–682

31. Gupta J, Singh V, Kumar I (2021) Florence-a health care chatbot. In: 2021 7th International conference on advanced computing and communication systems (ICACCS), vol 1. IEEE, pp 504–508
32. Zhang X, Wu J, He Z, Liu X, Su Y (2018) Medical exam question answering with large-scale reading comprehension. In: Proceedings of the AAAI conference on artificial intelligence, vol 32, no 1
33. Dalal S, Jain S, Dave M (2019) A systematic review of smart mental healthcare. In: Proceedings of the 5th international conference on cyber security & privacy in communication networks (ICCS)
34. Omoregbe NAI, Ndaman IO, Misra S, Abayomi-Alli OO, Damaševičius R (2020) Text messaging-based medical diagnosis using natural language processing and fuzzy logic. J Healthc Eng 2020
35. Bao Q, Ni L, Liu J (2020) HHH: an online medical chatbot system based on knowledge graph and hierarchical bi-directional attention. In: Proceedings of the Australasian computer science week Multiconference, pp 1–10
36. Harilal N, Shah R, Sharma S, Bhutani V (2020) CARO: an empathetic health conversational chatbot for people with major depression. In: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, pp 349–350
37. Cao YG, Liu F, Simpson P, Antieau L, Bennett A, Cimino JJ, Ely J, Yu H (2011) AskHERMES: an online question answering system for complex clinical questions. J Biomed Inf 44(2):277–288
38. Wu C, Luo G, Guo C, Ren Y, Zheng A, Yang C (2020) An attention-based multi-task model for named entity recognition and intent analysis of Chinese online medical questions. J Biomed Inform 108:103511
39. Ni Y, Zhu H, Cai P, Zhang L, Qui Z, Cao F (2012) CliniQA: highly reliable clinical question answering system. In: Quality of life through quality of information. IOS Press, pp 215–219
40. Oh K-J, Lee D, Ko B, Choi H-J (2017) A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In: 2017 18th IEEE international conference on mobile data management (MDM). IEEE, pp 371–375
41. Zhou H, Lei B, Liu Z, Liu Z (2019) Dut-bim at mediqa 2019: utilizing transformer network and medical domain-specific contextualized representations for question answering. In: Proceedings of the 18th BioNLP workshop and shared task, pp 446–452
42. Ren F, Zhou Y (2020) Cgmvqa: a new classification and generative model for medical visual question answering. IEEE Access 8:50626–50636
43. Soni S, Roberts K (2020) Evaluation of dataset selection for pre-training and fine-tuning transformer language models for clinical question answering. In: Proceedings of the 12th language resources and evaluation conference, pp 5532–5538
44. Alzubi JA, Jain R, Singh A, Parwekar P, Gupta M (2021) COBERT: COVID-19 question answering system using BERT. Arab J Sci Eng 1–11
45. Yadav S, Gupta D, Abacha AB, Demner-Fushman D (2022) Question-aware transformer models for consumer health question summarization. J Biomed Inf (2022):104040
46. Kadariya D, Venkataramanan R, Yip HY, Kalra M, Thirunarayanan K, Sheth A (2019) KBot: knowledge-enabled personalized chatbot for asthma self-management. In: 2019 IEEE international conference on smart computing (SMARTCOMP). IEEE, pp 138–143
47. Moon S, He H, Liu H, Fan JW (2022) RxWhyQA: a clinical question-answering dataset with the challenge of multi-answer questions. arXiv preprint arXiv:2201.02517

# An Analysis of Word Sense Disambiguation (WSD)

**Preethi Nanjundan and Eappen Zachariah Mathews**

**Abstract** Word sense disambiguation (WSD) is the method of using computer algorithms to determine the sense of arguments in the background. As a result of its difficult nature, WSD has measured an AI-complete problem, i.e., a problem whose key is as minimum as difficult as those posed by artificial intelligence. This article describes the task and introduces motives to resolve the ambiguity of words discussed throughout the text. This article summarizes supervised, unsupervised, and knowledge-based solutions. Senseval/semeval campaigns are described in relation to the assessment of WSDs, with the aim of an unbiased assessment of schemes working on numerous disambiguation errands. Finally, future directions, requests, open difficulties, and open problems are discoursed.

**Keywords** Word sense disambiguation · Word sense discrimination · WSD · Lexical semantics · Lexical ambiguity · Sense annotation · Semantic annotation

## 1 Introduction

In natural language processing (NLP), disambiguating different types of word senses (WSD) is a major challenge [1] that has quickly increased position since the advent of chatbots. When using WSD, homographs are differentiated based on their context words—identically spelled words that have different meanings and meanings in each of their surrounding sentences. WSD is the main NLP responsibilities that revolve about flawless solutions to sense arrangement and indication [2], and it often finds its way into applications surrounding NLP [3]. Numerous supervised methods toward the WSD unruly rely on models for training using sense-driven information [4]. Many of them, however, were not understandable. An example of NLP in a chatbot

---

P. Nanjundan (✉) · E. Z. Mathews
Department of Data Science, Christ University, Pune, Lavasa, India
e-mail: preethi.n@christuniversity.in

E. Z. Mathews
e-mail: eappen.mathews@science.christuniversity.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_22

is that it is used for speech recognition. The problem of chatbots' inability to distinguish between words with multiple meanings, distinct in different contexts, is well known from their real-life operation. In the sentence, "I want to buy a ticket for the upcoming movie, for instance, think of the term "book." You can order the term "book" as "reservation," but it is more like "reading material." It fails to explain how it reached that conclusion. To date, state-of-the-art NLP techniques have not succeeded in interpretability which is enhanced while classification is maintained accuracy. This paper is organized in a sequential order. A full description of the task is discussed in Sect. 2. The work in this section is concluded in Sect. 3.

## 2 Description of the Task

Disambiguation of word senses is the skill to control, using computational methods, which sense a word assumes from its usage in a specific setting. WSD is typically applied to one or additional manuscripts (while bags of words, i.e., sequences of arguments, may be utilized if the terms are naturally occurring). Despite the punctuation in the text, we can view it as being composed of arguments ($w_1$, $w_2$, …, $w_n$), and WSD can be conceptualized as transfer suitable senses or sense map(s) to the words in $T$, identifying the mapping between words and senses, in the sense that $A(i) \subseteq$ Senses $D(wi)$, wherever Senses $D(wi)$ is the usual of perceptions prearranged in a lexicon $D$ for word $wi$, 1 and $A(i)$ corresponds to those senses that are suitable in the setting of term $wi * T$. When plotting $A$, there can be additional sense assigned for individual word $wi \in T$, but naturally, only $A(i) = 1$ is given to each word. Natural language processing is likewise concerned with the classification [5], tagging parts of speech (redirecting parts of speech toward context-related target items), considering named entity resolution (utilizing predefined categories to classify specific texts), text categorization (i.e., assigning tags to manuscripts), etc. As a result, WSD is really composed of n separate organization responsibilities, where n is the scope of the lexicon. The generic task can be divided into two modifications:

(i) **A lexical sample** (or targeted word sense disambiguation): Disambiguation occurs when a system needs to distinguish between a limited conventional of mark arguments, typically happening one at a time.

(ii) **Word-by-word disambiguation:** Words in a text can be identified as nouns, adjectives, verbs, and adverbs by identifying their position in the text. As we approach the four main elements of WSD, we will look at how to select the word senses (i.e., classes), utilize outside information sources, represent setting, and choose an involuntary organization approach to use.

## *2.1 Choice of Word Senses*

The term sense refers to an agreed upon understanding of a word. For example, you may interpret the following verdicts:

(a) The mouse ate some cheese.
(b) Double-click the mouse to make changes.

The word mouse is cast off in the overhead sentences with two dissimilar senses: a minor long-tailed rodent (a) and a computer device (b). The two senses are obviously linked, as they may be request an equal thing; though, the thing's future usages are dissimilar. The instances brand it strong that causal the intelligence list of a term may be an important problematic in disambiguation of word sense: Are we future to allocate dissimilar lessons to the two amounts of the mouse in sentences (a) and (b)? Consider the sense list for the noun mouse, for example. (a) Should we add a more sense to the list for "a small long-tailed rodent" or does the primary sense include this intelligence? Because of such doubts, dissimilar selections are successful to be complete in numerous lexicons.

## *2.2 Knowledge Resources from Outside the Organization*

The principles of WSD are based on knowledge. Computer-assisted dictionaries, thesauri, glossaries, ontologies, etc., are also included in this research. Information on these and other matters can be found in Litkowski [6] and Agirre [7].

Assets are arranged as a structure:

(i) **These materials contain thesauruses**, which offer data about relations among arguments, including synonymity (like car, a substitute of motorcar), antonymy (as opposed to beautiful), and potentially others [8]. WSD very generally uses Roget's International Thesaurus [9].
(ii) **Machine-readable dictionaries (MRDs)**—Recent decades have seen the development of highly valuable sources of information for NLP, such as Collins English Dictionary and Oxford Dictionary of English, as well as Longman Dictionary of Contemporary English. To summarize the extensively explored LDOCE, WordNet's prolixity has been the most frequently used machine-readable lexicon in the NLP exploration community.
(iii) **Ontologies** which are conceptualizations of certain fields of curiosity are frequently comprising classification and a usual of semantic relations in this regard, in addition to rearrange and postulate WordNet, the SUMO upper ontology,
    (a) **Corpora**—Corpora are groups of documents used to learn language models. There are two types of corpora: sense-annotated and raw (i.e., unlabeled).

(b) **Raw corpora—**The Wall Street Journal (WSJ) corpus [10] is covering about 30 million arguments, and the Gig word corpus is consisting of 2 billion arguments of paper text [11].

(c) **The main and widely used sense-tagged corpus**—The MultiSemCor multisense corpus of English and Italian words, the SemCor corpus containing 4000 sense-tagged illustrations of nouns, adjectives, and verbs, and the Open Mind Word Expert dataset [12] were also utilized.

(d) **Samples contain the Word Draft Engine**—Web1T corpus [13] is a huge collection of manuscript co-occurrences that has quickly added approval in the WSD communal. One trillion words of the Web are used to generate frequency data for sequences of up to five words.

(e) **The second category of resources includes word frequency lists**, stop lists (a list of words without discrimination such as a, an, the, etc.), domain tags [14], etc.

The subsequent are some of the knowledge sources broadly used in the ground: WordNet.

**WordNet**. Synsets (sets of synonyms) based on psycholinguistic principles encode concepts (Miller et al. 1990; Fellbaum 1998). There are more than 117,000 synsets and 155,000 words in WordNet 3.0. As an example, consider the synset for the word automobile (remember superscripts and subscripts represent sense identifiers and parts of speech, respectively):

$$\left\{\text{lion}_n^1, \text{king of beasts}_n^1, \text{Panthera leo}_n^1\right\}.$$

Synsets can be thought of as sets of word senses that express (roughly) the same meaning. As described in Segment 2.1, the subsequent function assigns, respectively, part-of-speech tagged word, a WordNet sense corresponding to its aspect:

$$\text{Senses}_{W\,N} : L \times \text{POS} \rightarrow 2^{\text{SYNSETS}},$$

where SYNSETS is the complete set of synsets in WordNet. For instance:

$$\begin{aligned}\text{Senses}_{W\,N}(\text{lion}_n) = \Big\{&\left\{\text{lion}_n^1, \text{king of beasts}_n^1, \text{Panthera leo}_n^1\right\},\\ &\left\{\text{lion}_n^2, \text{social lion}_n^2\right\}, \left\{\text{lion}_n^3, \text{Leo}_n^3\right\},\\ &\left\{\text{lion}_n^3, \text{Leo}_n^4, \text{Leo the Lion}_n^4\right\}\Big\}.\end{aligned}$$

Every word sense can be identified unambiguously as belonging to a single synset. In particular, the synset of animal1, bird 1, canary 1, fish 4, and shark is unambiguously determined given animal$^1{}_n$. A WordNet semantic network excerpt showing animal$^1{}_n$ synset is shown in Fig. 1. The following information is provided by WordNet for each synset:

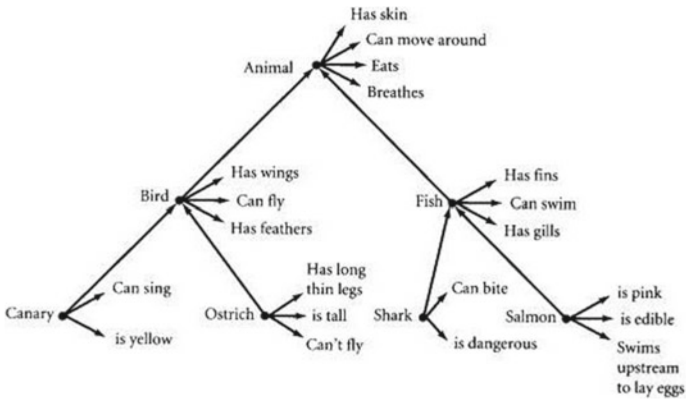


**Fig. 1** Extract of the WordNet semantic network

(a) This gloss is a documented definition of the synset perhaps including instances (e.g., a gloss of animal$^{1}_{n}$ would read "a living being equipped with skin, that has movement, eats, breathes, is loyal.").

(b) There are lexical and semantic connections between word senses and synsets. Lexical relatives denote word senses included in different synsets, whereas semantic relations characterize synsets in their entirety. The following are some examples of lexical relations:

(c) If X is antonymous with Y, then it means the opposite of Y (e.g., good$^{1}_{a}$ is the opposite of bad$^{1}_{a}$). Despite its name, a synonym exists for every word in the English language.

(d) Pertainymy: It is described by the adjective $X$ used to describe a noun (or a different word) $Y$ (e.g., dental$^{1}_{a}$ relates to tooth$^{1}_{n}$).

(e) $X$ nouns nominalize $Y$ verbs (e.g., service$^{2}_{n}$ nominalizes serve$^{4}_{v}$). Here are some semantic relations.

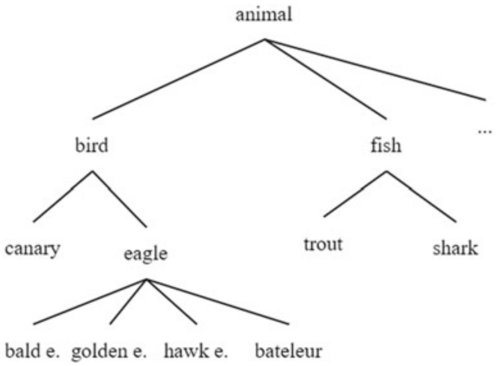
(f) An hypernym is a relation in which one (kind of) $X$ is the same as another (car$^{1}_{n}$ is a hypernym of motor vehicle$^{1}_{n}$). Nominal and verbal synsets exhibit hypernymy.

## *2.3 Contextual Representation*

In order to accomplish this, a preprocessing of the input text is typically performed, which can include the following steps (but not necessarily):

i. **Tokenization**, the process of dividing a text into tokens (usual words).
ii. **Part-of-speech tagging**, i.e., assigning grammatical categories to words ((Ram/NN, is/VBZ, a/DT, good/JJ, boy/NN, went/VBD, to/TO, school/NN), where DT, JJ, VBZ/VBD, and NN are tags for determinants, adjectives, predicates, and noun, respectively).

**Fig. 2** Extract of the WordNet domain tags' classification



iii. **Lemmatization**, which morphological variants are reduced to their most basic form (e.g., was → be, boys → boy).
iv. **Chunking**, the division of a document into pieces that are syntactically related (e.g., [Ram]NNP [went to school] VP are the noun phrase and the verb phrase of the example, respectively).
v. Analyzing syntactical structures of sentences is done by constructing syntax trees (Fig. 2).

Figure 3 shows an illustration of how the processing flow works. A preprocessing step can result in the representation of each word as a path of different structures, or in a more designed way, such as a tree or diagram that shows a relation between words. The context is signified by a set of features. These data include information resulting from preprocessing steps, including the parts of speech tags, the grammar relations, the lemmas, etc. The following features may be included in surveys:

(a) **Local features**: This type of characteristic describes the local framework of word usage, i.e., structures of a few adjacent words, such as word forms, parts of speech, relative positions to the target word, and so on.
(b) **Topical features**: Topical features, as opposed to local features, reflect larger circumstances (e.g., a gap of words, a phrase, a section, etc.) and are typically represented as bags of words.
(c) **Syntactic features**: Describe syntactic prompts and arguments as there is a relationship between the words in the sentence and the target term (keep in mind that these words are not always in the local context).
(d) **Semantic features**: A semantic feature represents aspects of a word such as its sense in context, a domain indicator, etc.

A feature vector can then be constructed based on individual word occurrence (regularly within a sentence). One of the following features is shown as a probable feature vector as illustrated in Table 1.

Consider Table 1, (a) the tank is full, and (b) the new tank has yet to be tested in the field, where tank is our vectors containing ten resident structures for the part-of-speech tags, and our target phrase is "The tank" which has two words on the left-hand

**Fig. 3** Sample of how text is preprocessed

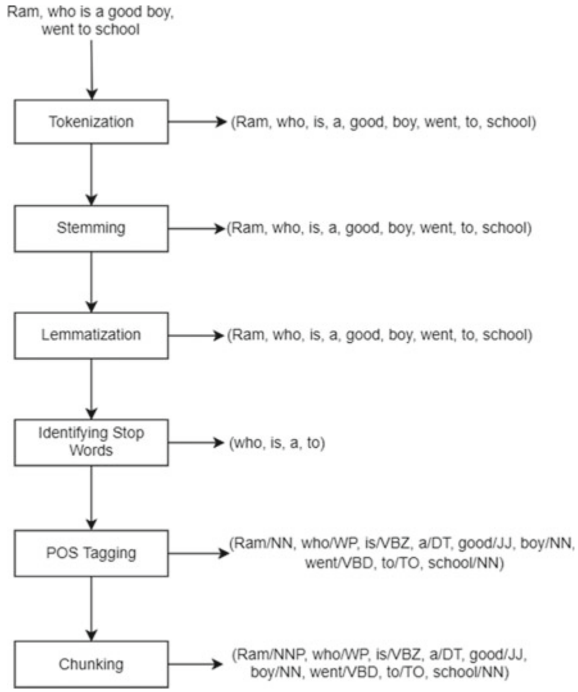


**Table 1** Nouns in sentences are represented by feature vectors

| Sentence | $w_{-2}$ | $w_{-1}$ | $w_{+1}$ | $w_{+2}$ | $w_{+3}$ | $w_{+4}$ | $w_{+5}$ | $w_{+6}$ | $w_{+7}$ | $w_{+8}$ | Sense tag |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (a) | – | Det | Verb | Adj | – | – | – | – | – | – | Vessel |
| (b) | Det | Adj | Verb | Adv | To | Verb | Verb | Prep | Det | Noun | Armored military vehicle |

and eight words on the correct, as well as a sense categorization tag (either VESSEL or ARMORED MILITARY VEHICLE in our illustration). Table 2 presents varying sized images for each word. A target word might be an n-gram (a sequence of n words combined with the target word), a bigram ($n = 2$), a trigram ($n = 3$), or a whole phrase or sentence.

Trees or graphs are frequently used as representations for word contexts that span the duration of a book. As training cases are frequently (but not always) conducted in this manner, flat depictions (such as background vectors) are best for supervised disambiguation approaches. The benefits of a structured portrayal lie in their use in both unsupervised and knowledge-based approaches, since they allow the full exploitation of the lexical and semantic links between ideas in computational lexicons and semantic networks.

**Table 2** Variations in word context sizes

| Context. size | Context. example |
|---|---|
| Unigram | …*Ram*… |
| Bigram | …*Ram* with…<br>…brought *Ram*… |
| Trigram | …brought *Ram* with…<br>…*Ram* with him…<br>…he brought *Ram*… |
| Window (size $\pm n$) $(2n + 1)$-grams | …and he brought *Ram* with him to… ($n = 3$)<br>…to school and he brought *Ram* with him to his home… ($n = 5$) |
| Sentence | Laxman went to school and he brought *Ram* with him to his home for lunch |
| Paragraph | It was Independence Day. Laxman went to school and he brought *Ram* with him to his home for lunch. After lunch, they played games and studied together for their exams. Both of them worked very hard and passed with flying colors |

## 3 Conclusion

Disambiguation of all words is incredibly helpful from a practical standpoint though we consider that disambiguating all content words is a bit academic. For example, in Senseval-3 all-words test set, tokens lemmatized as Bev make up roughly 8% of the total. This common verb does not appear to have much of an impact on the success of user inquiries in information retrieval systems. In vitro testing of WSD systems is not recommended because it reduces their performance unnecessarily and provides no information of benefits of end-to-end implementations. Several knowledge-based and supervised systems may perform with precision exceeding 90% and poor recall even after good sense distinctions are being utilized. This configuration may also have an impact on the concept of web semantics: the availability of technologies that can disambiguate would undoubtedly aid semantic interoperability. Disambiguation may be required only for a subset of a page content that conveys the resource's true content. Based on the meaning words convey, they can be disambiguated using computational lexicons and domain ontologies. Optimization ("disambiguate less, disambiguate better"), in both application-specific and relative settings, should be studied in subsequent assessment campaigns, in our opinion.

## References

1. Agirre E, Edmonds P (2007) Word sense disambiguation: algorithms and applications. Springer, Dordrecht
2. Navigli R, Camacho-Collados J, Raganato A (2017) Word sense disambiguation: a unified evaluation framework and empirical comparison. In: EACL
3. de Lacalle OL, Agirre E (2015) A methodology for word sense disambiguation at 90% based on large-scale crowdsourcing. In: SEM@NAACL-HLT

4. Liao K, Ye D, Xi Y (2010) Research on enterprise text knowledge classification based on knowledge schema. In: 2010 2nd IEEE international conference on information management and engineering, pp 452–456
5. Abney S, Light M (1999) Hiding a semantic class hierarchy in a Markov model. In: Proceedings of the ACL workshop on unsupervised learning in natural language processing (College Park, MD), pp 1–8
6. Litkowski KC (2005) Computational lexicons and dictionaries. In: Brown KR (ed) Encyclopedia of language and linguistics, 2nd edn. Elsevier Publishers, Oxford, U.K., pp 753–761
7. Ide N, Wilks Y (2006) Making sense about sense. In: Agirre E, Edmonds P (eds) Word sense disambiguation: algorithms and applications. Springer, New York, NY, pp 47–73
8. Kilgarriff A, Yallop C (2000) What's in a thesaurus? In: Proceedings of the 2nd conference on language resources and evaluation (LREC, Athens, Greece), pp 1371–1379
9. Roget PM (1911) Roget's international thesaurus, 1st edn. Cromwell, New York, NY
10. Charniak E, Blaheta D, Ge N, Hall K, Hale J, Johnson M (2000) Blip 1987–89 WSJ corpus release 1. Tech. rep. LDC2000T43. Linguistic Data Consortium (Philadelphia, PA)
11. Graff D (2003) English gig word. Tech. rep. LDC2003T05. Linguistic Data Consortium, Philadelphia, PA
12. Atkins S (1993) Tools for computer-aided corpus lexicography: the Hector project. Acta Linguistica Hungarica 41:5–72
13. Brants T, Franz A (2006) Web 1t 5-gram, ver. 1, ldc2006t13. Linguistic Data Consortium, Philadelphia, PA
14. Magnini B, Cavaglia‘ G (2000) Integrating subject field codes into WordNet. In: Proceedings of the 2nd conference on language resources and evaluation (LREC, Athens, Greece), pp 1413–1418
15. Bernard JRL (ed) (1986) Macquarie thesaurus. macquarie, Sydney, Australia

# Classification of Breast Invasive Ductal Carcinomas Using Histopathological Images Based on Deep Learning Techniques

**W. Jaisingh, N. Preethi, and S. Murali**

**Abstract** Women suffer from cancer, which is the main reason for death for females around the world. With the use of artificial intelligence, it is possible to predict and detect all types of cancers in the near future. It is not just women who can heal, and most breast cancers are caused by the most vulnerable type of breast. Eighty percent of all diagnoses of carcinoma are invasive ductal carcinomas (IDCs). In this paper, deep learning techniques are extended to support visible semantic evaluation of tumor areas, using convolutional neural networks (CNNs). A CNN is skilled ended a large number of photo covers (tissue areas) after Whole Slide Images (WSI) to study ranked part-based total image. About 600 normal image patches and 200 breast invasive ductal carcinomas are selected for the experiment. It was intended to amount classifier correctness in the detection of IDC tissue areas in Whole Slide Images. We achieved excellent measurable outcomes for an automated finding of IDC areas with our technique. The results are evaluated based on performance measures and compared with a different number of neurons, and the results are highlighted.

**Keywords** Breast cancer · Invasive ductal cancer · Histology microscopy image · Deep learning · Pattern recognition

---

W. Jaisingh (✉)
School of Information Science, Presidency University, Bangalore, Karnataka 560064, India
e-mail: jaisinghw@gmail.com

N. Preethi
Department of Data Science, Christ University, Bengaluru, India
e-mail: preethi.n@christuniversity.in

S. Murali
Department of Mathematics, Coimbatore Institute of Technology, Coimbatore 641014, India
e-mail: murali.s@cit.edu.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_23

## 1 Introduction

Segmentalizing semantic images simplifies the process of representing and interpreting images by grouping pixels that share a common definition to separate concepts. The detection and segmentation of tumors are essential steps in medical imaging for diagnosis and disease characterization. This is particularly true in histopathology, where tissue samples can contain a variety of cell types and numbers. Modern screening systems and new technologies have made it possible for breast cancer to be detected earlier. Estimation of the rate of ductal carcinoma in situ (DCIS). As a complication of invasive breast cancer, DCIS is common. An important prognostic factor has been the size of the DCIS associated with the tumor. Breast-slice surgery and breast-conserving surgery patients are in danger of local reappearance.

In the USA, malignancies are the leading cause of death as a result of medical issues affecting people worldwide [1]. In the USA, there are about 18.1 million new diseases analyzed every year, as per the International Agency for Research on Cancer (IARC). About 9.6 million disease cases and 9.6 million passing were affirmed in 2018 [2]. The diagram shows the worldwide frequency of bosom malignancy in females. The outline is shading coded by the sorts of disease. Malignant growth of the bosom is coded with the 'pink' shading. Worldwide, breast cancer is the most well-known disease among ladies. Throughout the planet, in the year 2018, there were assessed to be 20.8 lakh instances of bosom disease. In ladies, bosom malignancy represented 24.2% of all tumors. In India in 2018, 162,468 ladies were recently determined to have breast cancer and 87,090 ladies passed on from breast cancer [2–6]. Figure 1 shows the estimated number of incident cases and deaths.

As per Schnitt SJ et al., breast cancer with an extensive intraductal component (EIC) consists of more than 25% of the ductal component. DCIS may occur within or outside of the main tumor mass [8]. EIC is also related to an advanced proportion of optimistic test results. Surgeons are frustrated by re-excisions, margins, and local recurrences. With a strong EIC, surgeons are more likely to use more aggressive techniques [9, 10]. Without obtrusive sickness, high-grade DCIS is related to a higher danger of repeat [11].

DCIS is related to a high frequency of invasive ductal carcinoma (IDC) (25–80%, depending on the series). However, the relationship between these two bodies has not been thoroughly explored. A few studies have investigated the link between DCIS and intrusive bosom malignant growth as a prognostic factor; however, these investigations have demonstrated to be exceptionally questionable [12–14].

Conversely, [15] detailed a high pace of repeat when DCIS and IDC were corresponded [6]. Conny Vrieling et al. discovered that promoter after monitoring bosom a medical procedure diminished nearby repeat in high-hazard patients (50 years and related DCIS) and, all the more fundamentally, the impact of mediation in high-hazard patients (50 years and related DCIS).

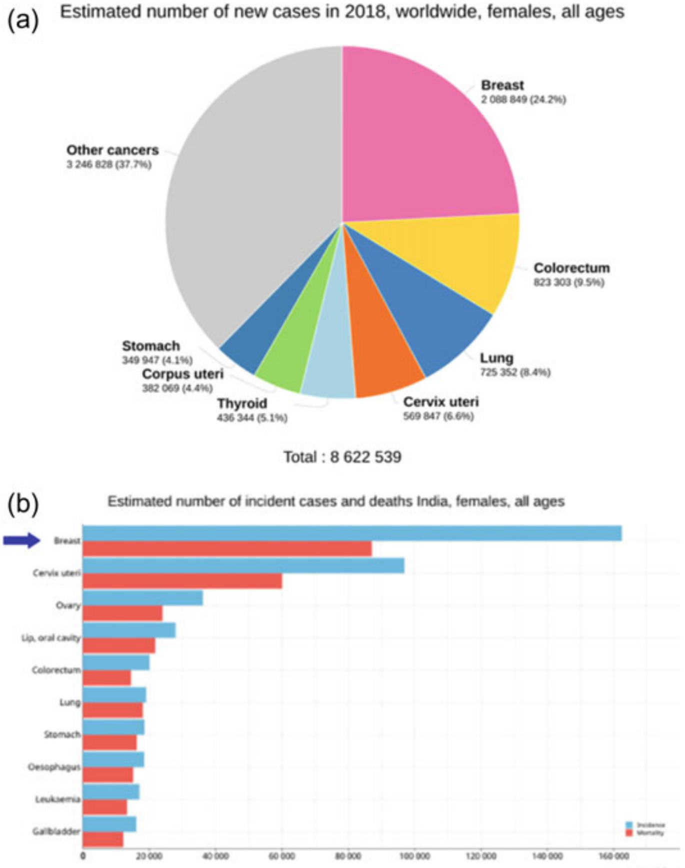


**Fig. 1** **a** Estimated number of cases in 2018, **b** estimated number of incident cases in India [7]

The following is a breakdown of the structure of this document: Sect. 2 contains a list of related works. As a key contribution of the research, the dataset is well-defined in Sect. 3. The planned methodology is presented in Sect. 4. The discussion of the faltering results and assumptions of the paper are presented in Sects. 5 and 6.

## 2 Related Work

The methods developed by different authors on breast cancer are described in this section.

Utilizing pictures from fine-needle biopsy microscopy, [16] fostered a PC-supported bosom malignancy analysis. Five hundred cytological pictures from 50 patients (10 pictures for each understanding) were remembered for this article. The patient is assessed and arranged into two threatening gatherings. The pictures were RGB in BMP design, with a pixeled tally of 704 578. K-means (KM), fuzzy C-means (FCM), and Gaussian were utilized to figure the pictures.

BreaKHis is a dataset for the recognition of breast cancer histopathologic images proposed by [17]. About 7909 images of breast cancer histopathology were provided by the authors. In total, 82 patients were seen. Magnification options include 40×, 80×, and 100×. Magnifications of 100×, 200×, and 400×, with binary classifications: malicious and gentle. As a starting point, identify patterns.

As a result of [18], a far-off PC-aided strategy to recognize and counteract breast cancer growth has been proposed. A total of 11,502 nuclei were analyzed from 92 breast cytological images. Images ranged in size from 640 to 960 pixels. From a wide range of pixels, a 480 × 2560 resolution can be achieved. To determine whether malignant cells or benign ones are present in this study, four machine learning algorithms were applied.

In their study [19], Yuankai K. Tao et al. suggested using nonlinear microscopy (NLM) to identify breast pathologies. A total of 179 samples were collected beginning with 50 patients utilizing the NLM. The picture dataset included ordinary bosom tissue, fibroadenoma, far-reaching ductal hyperplasia, and different anomalies. Lobular carcinoma in situ (LCIS), intrusive lobular carcinoma, and fibrocystic changes are two kinds of malignancy.

Based on the above literature review, we have examined various methodologies for detecting invasive ductal carcinoma in histopathological microscopy images.

In the rest of the paper, the approach is described in Sector 3 and the experimental assessment process and the dataset are defined in Sector 4. In Sector 5, the results and discussion are presented. In Sector 6, the key findings and future directions are presented.

## 3 Methodology

A novel CNN classification technique is used in this work with relu, max_pooling, soft_max, and sigmoid function. The steps of the framework are as follows.

### *3.1 Convolutional Neural Networks*

CNN programs apply neighborhood characteristic detectors or filters to the entire photograph to grade the communication among character photograph covers and cross-styles in the instruction usual. A combination or merging feature is then implemented to decrease the dimensionality of the characteristic space. The photograph covers from Stage 1 is then fed into a three-layer CNN structure in which layers are utilized for convolution and pooling at the same time, and the last layer is completely associated.

CNN design comprises the accompanying layers: convolutional layer, corrected direct unit (ReLU) enactment layer, group standardization layer, pooling layer, dropout layer, completely associated layer, and softmax layer [20–22]. Our proposed engineering consists of five convolutional layers, followed by an entirely associated layer and a softmax layer for arrangement.

Accordingly, the convolutional layer will decide the yield of neurons that are associated with neighborhood locales of the contribution through the computation of the internal item between their loads [20]. Channels of differing sizes (for example, 3 * 3, 5 * 5, 7 * 7) are convolved with the info information to extricate various highlights. Each convolutional layer is made out of three convolution channels, a ReLU, a bunch standardization layer, and the greatest pooling layer. We utilize the 3 * 3 convolutional channel to dependably catch the IDC usefulness. Moreover, the convolution layer has similar support, and the yield is a similar size to the information. In Table 2, you can see the number of convolutional channels remembered for each channel. Sort things out. Convolution yield is passed to the ReLU layer.

The pooling layer will essentially do a down test by squaring up the spatial dimension of the information to reduce the number of boundaries in the enactment [20]. Eventually, the model's computational complexity will decrease, the portrayal will be smaller, and fewer boundaries will be used. CNNs use a layer of pooling that has a maximum size and scope of 2 * 2.

The last convolutional layer, for example, the fifth convolutional layer, is a dropout layer. Dropout might be utilized to take care of overfitting issues. During preparation, irregular units (and their associations) are dropped from the neural organization [22]. This keeps units from co-adjusting an unnecessary measure of dropout and fundamentally lessens overfitting contrasted with other regularization techniques [21, 22]. For the dropout layer, we use an eliminate likelihood of 0.25.

Neurons in the completely associated layer are directly wired to neurons in two neighboring layers, but they are not officially associated with either of them. The neurons in each layer have 128 units each. The neurons in the softmax layer have two units each. Softmax is completely applied to the yield values. Likelihood dissemination is made by normalizing the yield esteems with the softmax work.

The remarkable capacity is applied to each yield esteem obtained by CNN from the positive class neurons of categorized areas to generate a value among 0 and 1, which can be interpreted as a probability. The original WSI coordinates and probabilities for each patch are then stitched together to generate the IDC likelihood plot over WSI.

## 4 Materials and Methods

### 4.1 Description of the Dataset

The information associate comprises digitized breast cancer (BC) histopathology slides of 980 images obtained from the Kaggle dataset. Slides were examined entire slide at 40× amplification (0.25 m/pixel goal) with an entire slide scanner. Regardless of their huge size (on the request for 1010 pixels), working on entire slides of histopathology photographs is unimaginable. This investigation down inspected each WSI (by a factor of 16:1) to a goal of 4 m/pixel.

### 4.2 Annotate the Ground Truth

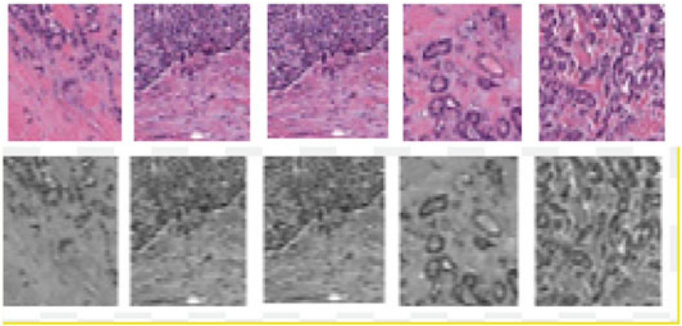
To set up the ground truth for IDC regions, a pathologist physically outlined the malignant growth zone utilizing computerized histopathology slides. The comments were made with Aperio's ImageScope, a free survey program. Since proficient ground truth explanations require a lot of time and exertion, this examination centers around generally coarse comments at the amplification of 2× or less. Thus, some stromal and non-obtrusive tissue was incorporated close to the proposed IDC tissue areas.

### 4.3 Building Datasets Based on Image Patches

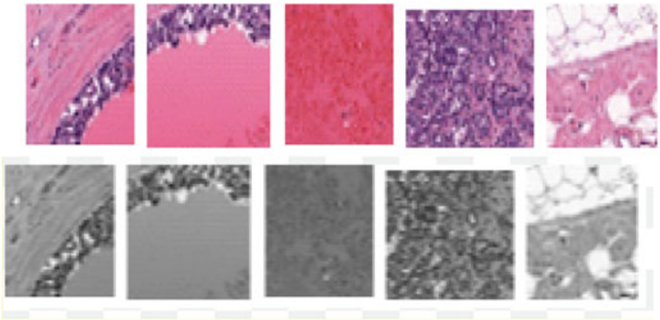
The original dataset was composed of 162 whole mount slide images scanned at 40x, each containing a breast cancer (BCa) specimen. In total, two hundred and seventy-five thousand patches (198,738 IDC negatives and 78,786 IDC positives) were extracted. U is the patient ID (10253idx5), X is the x-coordinate of where this patch was cropped from, Y is the y-coordinate of where this patch was cropped from, and C is the class, where 0 is non-IDC and 1 is ID.

To grade the communication between character photograph covers and name styles in the instruction place, CNC programs apply neighborhood characteristic detectors or filters to the entire photograph. To reduce the dimensionality of the characteristic space, an aggregation or pooling feature is implemented. The patches from Step 1 are fed into a three-layer CNN structure in which layers are used for convolution and pooling simultaneously, with the final layer being fully connected.

A CNN design contains the accompanying layers: convolutional layer, amended direct unit (ReLU) initiation layer, cluster standardization layer, pooling layer, dropout layer, completely associated layer, and softmax layer [20–22]. Figure 2 shows the examples of positive and negative samples of image patches.



a) Positive examples



b) Negative examples

**Fig. 2** Examples of image patches

## 5 Result and Discussion of the Experimental Work

Our CNN-based framework is thought about in contrast to a bunch of carefully assembled highlights that are cutting edge (worldwide highlights or histopathology highlights). The following are the subtleties of the two methodologies.

In our framework, convolutional pooling layers are just like the completely associated layer, and our framework utilizes a three-layer CNN design with 16, 32, and 128 neurons, separately. In all preliminaries, a fixed convolutional part of size 8 × 8 and a pool portion of size 2 × 2 were utilized.

The entire dataset is divided into training and testing sets. The validation dataset consists of 150 normal and 150 abnormal images from Kaggle's dataset. The training set contains 600 normal images and 200 invasive ductal carcinomas. Anaconda and Keras framework are used to train and test the classification CNN method.

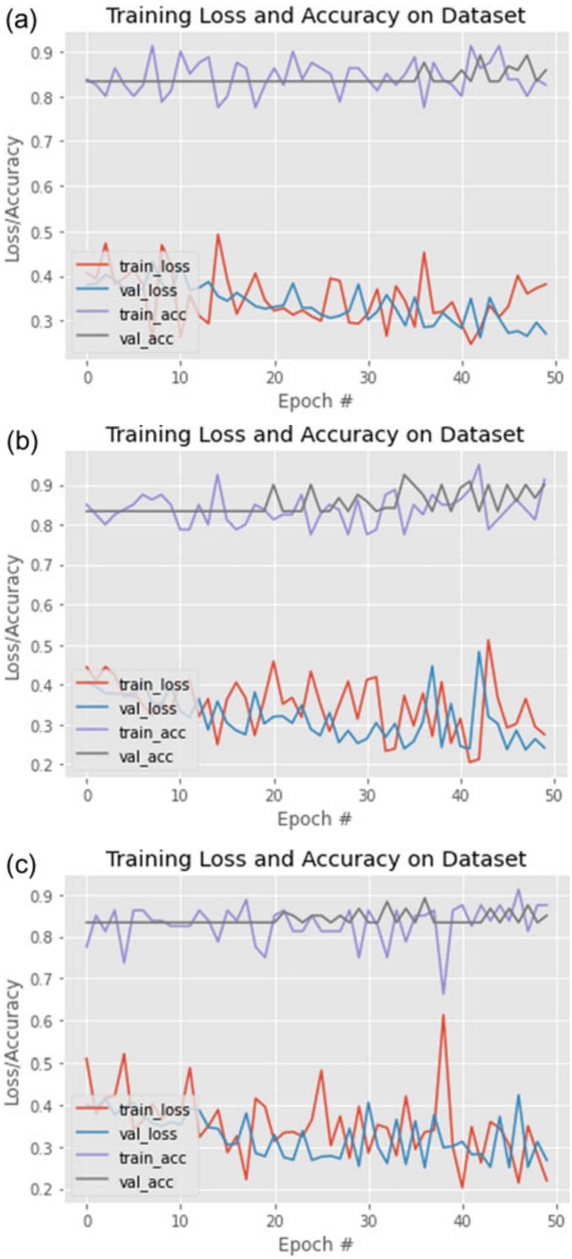
Training and testing sets are processed with a cross-entropy loss function. In order to increase the accuracy of the model, training is done for between 3 and 50 times. The evaluation parameters are summarized in Table 1. Using validation data, it has been observed that 89% accuracy has been reached.

The number of epochs during the training period was increased to combat the underfitting problem. The experiment is executed by using CNN with 32, 64, and 128 neurons. In Table 1, it is observed that CNN with 32 neurons has got 89.83% of validation accuracy, 64 neurons have an accuracy of 92.16%, and 128 neurons have an accuracy of 88.67%. From Table 1 and Fig. 3, it is observed that CNN with 64 neurons achieves an accuracy of 92.16%.

**Table 1** Overview of training loss, training accuracy, validation loss, and validation accuracy

| Epochs | Training loss | Training accuracy | Validation loss | Validation accuracy |
|---|---|---|---|---|
| *32 neurons* | | | | |
| 1 | 0.4064 | 0.8375 | 0.3796 | 0.8333 |
| 2 | 0.3929 | 0.8250 | 0.3822 | 0.8333 |
| 3 | 0.4724 | 0.8000 | 0.4030 | 0.8333 |
| 4 | 0.3821 | 0.8625 | 0.3911 | 0.8333 |
| 5 | 0.3954 | 0.8250 | 0.3715 | 0.8333 |
| 45 | 0.3077 | 0.9125 | 0.3057 | 0.8333 |
| 46 | 0.3318 | 0.8375 | 0.2723 | 0.8667 |
| 47 | 0.4007 | 0.8375 | 0.2766 | 0.8583 |
| 48 | 0.3604 | 0.8000 | 0.2657 | 0.8917 |
| 49 | 0.3725 | 0.8375 | 0.2959 | 0.8933 |
| 50 | 0.3813 | 0.8250 | 0.2710 | 0.8983 |
| *64 neurons* | | | | |
| 45 | 0.3677 | 0.8225 | 0.3022 | 0.8942 |
| 46 | 0.2921 | 0.8375 | 0.2376 | 0.9012 |
| 47 | 0.3019 | 0.8725 | 0.2550 | 0.8956 |
| 48 | 0.3645 | 0.8375 | 0.2368 | 0.9015 |
| 49 | 0.2749 | 0.9125 | 0.2437 | 0.9167 |
| 50 | 0.2551 | 0.9212 | 0.2218 | 0.9216 |
| *128 neurons* | | | | |
| 45 | 0.3960 | 0.7500 | 0.2335 | 0.8667 |
| 46 | 0.2851 | 0.8625 | 0.4051 | 0.8332 |
| 47 | 0.3522 | 0.8250 | 0.3003 | 0.8432 |
| 48 | 0.3449 | 0.7865 | 0.2604 | 0.8833 |
| 49 | 0.2788 | 0.8625 | 0.3647 | 0.8543 |
| 50 | 0.3201 | 0.8175 | 0.2379 | 0.8867 |

**Fig. 3** Estimate of each epoch’s loss and accuracy curves, **a** CNN with 32 neurons, **b** CNN with 64 neurons, **c** CNN with 128 neurons



## 6 Conclusion

The discovery of women having intrusive ductal carcinoma was an incredibly intriguing and groundbreaking discovery on bosom cancer. As the first step in diagnosing and treating BCa, accurate and repeatable IDC detection is vital. In previous research on histopathology tumor diagnosis, researchers have sought to resolve this challenge by integrating machine learning and handmade features. A CNN classifier with 32, 64, and 128 neurons has implemented in the experiment. Using the results of the experiments, it is found that effective CNN with 64 neurons gives a good classification accuracy compared to 32 and 128 neurons.

## References

1. Rebecca L, Seigel MPH, Kimberly D, Miller MPH, Ahmedin Jemal DVM (2019) Cancer statistics. CA A Cancer J Clin 69(1):7–34
2. Wild CP, Weiderpass E, Stewart BW World cancer report: cancer research for cancer prevention. In: World cancer reports. World Health Organization, IACR, ISBN: 978-92-932-0448-0
3. Institute for Health Metrics and Evaluation (IHME) (2018) Global burden of disease study 2017. GBD 2017
4. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchid O, Dandona R, Dandona L et al (2017) Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. JAMA Oncol 3(4):524–548
5. World Cancer Research Fund (2018) Breast cancer statistics
6. DeSantis C, Siegel R, Bandi P, Jemal A (2011) Breast cancer statistics. CA A Cancer J Clin 61(6):408–418
7. http://www.breastcancerindia.net/statistics/stat_global.html
8. Schnitt SJ, Connolly JL, Harris JR et al (1984) Pathologic predictors of early local recurrence in Stage I and II breast cancer treated by primary radiation therapy. Cancer 53:1049–1057
9. van Deurzen CH (2016) Predictors of surgical margin following breast-conserving surgery: a large population-based cohort study. Ann Surg Oncol 23:627–633
10. Elling D, Vesper AS, Fiedler B et al (2001) Intraductal component in invasive breast cancer: analysis of 250 resected surgical specimens. Breast 10:405–410
11. Bijker N, Petersen JL, Duchateau L et al (2001) Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. J Clin Oncol 19:2263–2271
12. Chagpar AB, McMasters KM, Sahoo S, Edwards MJ (2009) Does ductal carcinoma in situ accompanying invasive carcinoma affect prognosis? Surgery 146:561–567, discussion 7–8
13. Wong H, Lau S, Yau T et al (2010) Presence of an in situ component is associated with reduced biological aggressiveness of size-matched invasive breast cancer. Br J Canc 102:1391–1396
14. Sebastian Sebastian C, Garcia Mur C, Cruz Ciria S et al (2016) Imaging and histologic prognostic factors in triple-negative breast cancer and carcinoma in situ as a prognostic factor. Radiologia 58:283–293
15. https://pubmed.ncbi.nlm.nih.gov/9701363/
16. Kowal M et al (2013) Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Comput BiolMed October 43(10):1563–1572
17. Spanhol Fabio A, Oliveira Luiz S. Caroline Petitjean and Laurent Heutte, “A dataset for breast cancer histopathological image classification. In: IEEE Transaction on biomedical engineering; 2015. vol 00, No 00.

18. Mourice George Y et al (2014) Remote computer-aided breast cancer detection and diagnosis system based on cryptological images. IEEE Syst 8(3)
19. Tao Yuankai K et al (2014) Assessment of breast pathologies using nonlinear microscopy. Proc Natl Acad Sci Unit States Am 111(43):15304–15309
20. O'Shea K, Nash R (2015) An introduction to convolutional neural networks. ArXiv e-prints
21. Bjorck J, Gomes C, Selman B (2018) Understanding batch normalization. NeurIPS
22. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15:1929–1958

# Heart Disease Detection Using Machine Learning Techniques

**Satabdi Swain, Sujata Chakravarty, Bijay Paikaray, and Harshvardhan Bhoyar**

**Abstract** Technology is growing rapidly and entering into almost all fields including health care. Coronary heart diseases (CHDs) are accounting for the larger number of deaths, which are estimated to be 17 million, i.e., 34% of the global death. Heart disease is considered to be one of the riskiest diseases, yet considered the most manageable disease. Hence, if the disease could be predicted early, it could be more beneficial. Here, 'Artificial Intelligence (AI) and Machine Learning (ML)' come into action as data science contributes a significant role in the field of healthcare, as it helps in analyzing and processing huge data. This paper explains how based on medical data and habits of a person, an ML model can predict whether a person will be a risk of CHD in the next 10 years or not. This experiment tends to find out the most accurate model for heart disease prediction by using various classification techniques of ML such as artificial neural networks (ANNs), logistic regression (LR), random forest (RF), and decision tree (DT). It is thus found that RF has the highest accuracy.

**Keywords** Logistic regression · Coronary heart disease · Random forest · Risk factor · Machine learning · Decision tree · Artificial intelligence · Neural network

## 1 Introduction

As we all know that heart is the most crucial and delicate organ of our body, which plays the most important role in all living organisms, the heart pumps blood to other organs and helps in purifying the blood. A small malfunction of the heart can cause deadlier problems and discomfort in other parts of the body as well [1]. Generally, CHD is the situation when an inadequate amount of blood is not supplied to the

---

S. Swain · S. Chakravarty
Department of CSE, Centurion University of Technology and Management, Jatani, Odisha, India

B. Paikaray (✉)
School of Information and Communication Technology, Medhavi Skills University, Sikkim, India
e-mail: bijaypaikaray87@gmail.com

H. Bhoyar
Sri Sri University, Cuttack, Odisha, India

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_24

heart because of the deposit of cholesterol or fats in the coronary arteries. Smoking, high blood pressure, high cholesterol, or diabetes are the main reasons for CHD. If these data are analyzed with high precision, then the possibility of having CHD in the next 10 years can be predicted. For this, we require a model with maximum accuracy because even a tiny mistake can cost a life. In this experiment, we have designed classification models using various ML algorithms like 'LR', 'RF', 'DT', and 'ANN'. We have done binary classification where the outcome is either '0' or '1', where 0 means no risk and 1 means yes of heart disease, and there is a risk. We will compare the accuracy of each model and find the best out of it.

Several techniques are used for detecting the presence or risk of CHD, by a physician like angiography, but there are certain limitations as well as they are very expensive, painful, risky, and time taking. But detecting the disease using ML algorithms is quite easier and quick, the only thing is that it needs to be accurate.

As we are dealing with the healthcare domain, it requires a proficient and enormous amount of data to train the machine learning model properly [2]. The collected data are also very noisy and too immense to be grasped by human minds [3]. So, the data need to be cleaned and preprocessed before going for model fitting and training. The extraction is done by data mining of important information from these bulkier and noisy datasets that results in predicting the data using classification techniques [4].

The rest of the paper is organized in the following manner. In Sect. 2, literature survey is discussed. Section 3 represents difficult machine learning algorithms used. In Sect. 4, information about the dataset is provided. Section 5 has the data preprocessing steps. In Sect. 6, accuracy and results are discussed. Finally, Sect. 7 concludes the paper and focuses on future work.

## 2 Literature Survey

Heart disease detection using machine learning techniques is an in-progress work for the past 20 years. Most of the papers have shown works related to this disease using various classification techniques of machine learning such as Naive Bayes, logistic regression, decision tree, support vector classifier, neural networks [5]. On various data of patients from across the world, papers differ based on the selection of different parameters on which the methods are applied. Different authors have applied different methods, machine learning models, parameters, and different datasets to test the accuracy of the models. It can be seen that each algorithm has the strength to register defined algorithms [6]. Many authors claim decision trees to be the most considerable to aid in the early diagnosis of heart illness. In [7], support vector models, decision trees, K-nearest neighbor, and Naive Bayes are used for the experiment, but it shows how Naive Bayes outperforms other models. It clearly explains the working and mechanism of the mentioned algorithms. In [8], the multilayer perceptron technique of ANN is implemented for obtaining better accuracy. Information about different mobile applications like my heart, Fitbit, and health gear is manifested in this paper. Edges and drawbacks of the different machine learning algorithms can be known from [9]. Paper [10] provides an insight into existing algorithms like ANN, DT, Naive Bayes, fuzzy logic, SVM, and K-nearest neighbor. They examined a wide range of health-related sensors and technology. They discovered certain problems that needed to be looked at. Chiuchisan and Geman conceptualized the house observation system and decision support system in the year 2014 [11]. This approach helped the persons who are having Parkinson's disease get home monitoring, diagnoses, prescriptions, medical treatment, rehabilitation, and development. Over the past ten years, the introduction of wireless technology in the health care monitoring system (WHMS) has received a lot of attention from both the research as well as from the business community. It has been observed that weighted associative classifier performance and other machine learning classifier methods have improved the detection of cardiac anomalies [12]. In [13], a completely operational body area network system which is wireless (WBAN) was suggested. Medical bands were employed in the system's design to collect physiological data from sensors. To reduce the interference between the sensors and equipment, the author selected particular medical bands. A hospital doorway wireless board was utilized in this process to apply the multi-hopping approach and expand the operational range. A cardiac illness forecast system grounded on SEM and FCM was created by Manpreet Singh. The information from the Canadian Community Health Survey 2012 dataset was verified. They employed twenty important characteristics of SEM by which the weight matrix for the FCM model is created. It then made predictions about the likelihood of cardiovascular illnesses. The co-relation between CCC-121 and 20 qualities, where the variable CCC-121 shows whether the respondent has heart disease, forms an SEM model. Big data was used by Ghadge et al. [14] to do research on a smart cardiac arrest prediction system. Because heart attacks are so common, it is important to detect them quickly and accurately. Finding a model of a smart cardiac arrest detection system that makes use of big data and data mining modeling approaches was the major goal of this research study. This technology could extract undiscovered information on heart illness from a specific heart disease database.

## 3 Algorithms Used

### *3.1 Logistic Regression Methodology*

It is a supervised machine learning algorithm that can be used for binary class classification [15]. It is a statistical model which classifies the model based on the input values. It predicts the target value or the dependent variable based on independent variables. There are different types of logistic regression, but as the target variable in the data set used in this paper is binary, so binomial logistic regression is used. The equation for a multiple binary logistic regression is as follows:

$$\pi(x) = \frac{e^{x+\beta x}}{1 + e^{x+\beta x}} \tag{1}$$

where response is the presence or absence of characteristics, the predictor is the variable observed for $\pi(x)$, and $\beta$ is the bias.

### 3.2 Decision Tree Methodology

It is a supervised machine learning algorithm used for both classification and regression problems [16]. Some datasets are very complex, where we cannot divide the data with a single line, and the data are categorical. Here, decision tree algorithms are used which will divide the data again and again. There are three nodes in a decision tree, based on which the model analyzes and classifies the result [7]. The three nodes are the root node—it is the main node from where division starts, branches—here, decision rules are set according to which further nodes are divided, and leaf node—these are the end nodes, which has the result.

The formula for root node is as follows:

$$\text{Information Gain} = \text{Class Entropy} - \text{Entropy Attributes} \tag{2}$$

The formula for class entropy is as follows:

$$(Pi + Ni) = -\frac{P}{P+N}\log_2\frac{P}{P+N} - \frac{N}{P+N}\log_2\frac{N}{P+N} \tag{3}$$

Here, $P$ means the possibility of Yes, and No means the possibility of No.

For finding entropy attributes (Fig. 1):

$$\sum \frac{Pi + Ni}{P+N} \tag{4}$$

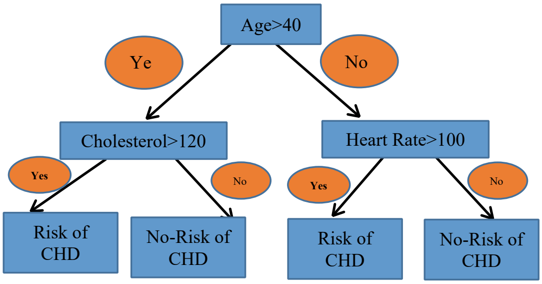


**Fig. 1** Example of decision tree

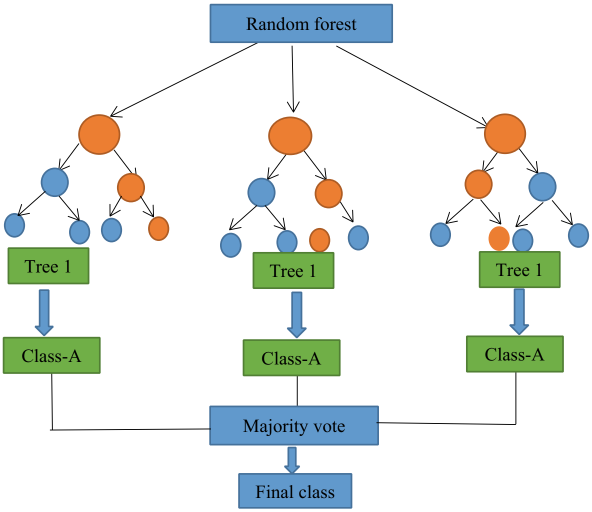


**Fig. 2** Architecture of random forest

## 3.3 Random Forest Methodology

Both classification and regression issues can be solved with this supervised machine learning method [17]. It is a classification model that contains several subtrees of a subset of given data and considers the average to enhance the accuracy. A random forest is made up of several decision trees. In the end, it evaluates all the trees and finds the tree with maximum votes and then classifies it into the final class. The expected accuracy of the model is comparatively as it is an amalgamation of several decision trees (Fig. 2).

## 3.4 Artificial Neural Network Methodology

Artificial neural networks resemble biological neural networks seen in human brains, but tough not similar [18]. It has three layers: input layers—input data are given, based on which model will work; hidden layer—the main layer where actual computation work is done; and an output layer—this layer has the final result (Fig. 3).

The ANN takes the inputs and computes the weighted sum of the inputs and creates a bias or a threshold value that is used for activation of the neurons.

Transfer function:

$$\sum_{i=1}^{n} Wi * Xi + b. \tag{5}$$

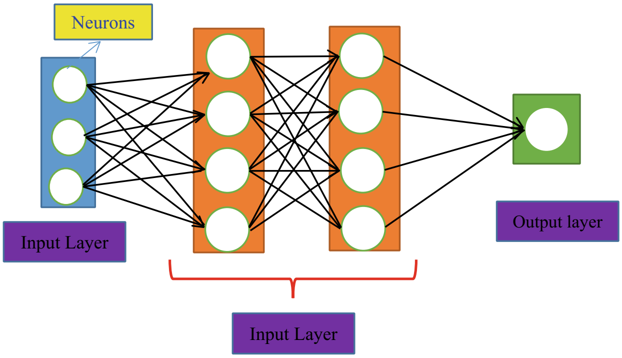


**Fig. 3** Architecture of artificial neural network

## 4 Dataset

In this study, we used online data which is downloaded from the Kaggle website. These data are from an ongoing survey of the people of Framingham, Massachusetts. Our dataset has data of 4000+ patients where there are 16 attributes which have information about them like age, gender, BP medications, etc. It has a binary label named TenYearCHD which is used as our target variable, where 0 means no heart disease and 1 means that there is a risk of heart disease.

### *4.1 Feature Description*

Table 1 gives a piece of detailed information about the features. Here, TenYearCHD is used as the target/dependent variable and others are used as independent variables. As the education column is not necessary for disease detection, it can be dropped from the dataset.

### *4.2 Attributes' Graph*

See Fig. 4.

### *4.3 Architecture Diagram*

See Fig. 5.

**Table 1** Piece of detailed information about the features

| Sl. No. | Feature | Description |
|---|---|---|
| | Male | It is a binary label where 0 means female and 1 means male |
| | Age | Specifies the age of the patient |
| | Education | Tells about the education of the patient |
| | Current smoker | Binary label where 1 means that the patient is a current smoker and 0 means that the patient does not smoke |
| | Cigarettes per day | It tells about how many cigarettes a person takes per day on average |
| | BP Meds | 1 means the person takes blood pressure medicine and 0 means do not take medicines |
| | Prevalent stroke | 1 indicates no prevalent strokes and 0 indicates had a stroke before |
| | Prevalent hyp | 1 indicates had high blood pressure before and 0 indicates no prevalent high BP |
| | Diabetes | 1 indicates that the patient has diabetes and 0 indicates no diabetes |
| | Total cholesterol | Gives information about the total cholesterol of the patient (normal:100–120) |
| | Systolic BP | Systolic blood pressure of the patient (normal:120–140) |
| | Diastolic BP | Diastolic blood pressure of the patient (normal:80–100) |
| | BMI | Body mass index |
| | Heart rate | Heart rate of the patient (normal:72–100) |
| | Glucose | The blood glucose level of the person |
| | TenYearCHD | 0 specifies no risk of coronary heart disease and 1 specifies that there is a risk of coronary heart disease in the next 10 years |

## 5 Data Preprocessing

As data related to the healthcare domain are too big and contain a lot of noise, so it needs to be cleaned and normalized for using it for classification. At first, we have to find the null values from the dataset and remove them to avoid a not a number (NaN) error. Then, feature extraction is done where the independent and dependent variables are identified and separated and the unnecessary features are removed. In this paper, the education variable is not required, so it is removed (Table 1). Then, the data need to be normalized to convert them within a range of 0 and 1. To normalize the data in this range, the following formula is used:

$$\boldsymbol{x}\mathbf{norm} = (\boldsymbol{xi} - \boldsymbol{x}\mathbf{min})/(\boldsymbol{x}\mathbf{max} - \boldsymbol{x}\mathbf{min}) \tag{6}$$

Here, **Xnorm is the** *i*th normalized value in the dataset, **Xi is the** *i*th value in the dataset, **Xmax is the** minimum value in the dataset, and **Xmin is the** maximum value in the dataset.

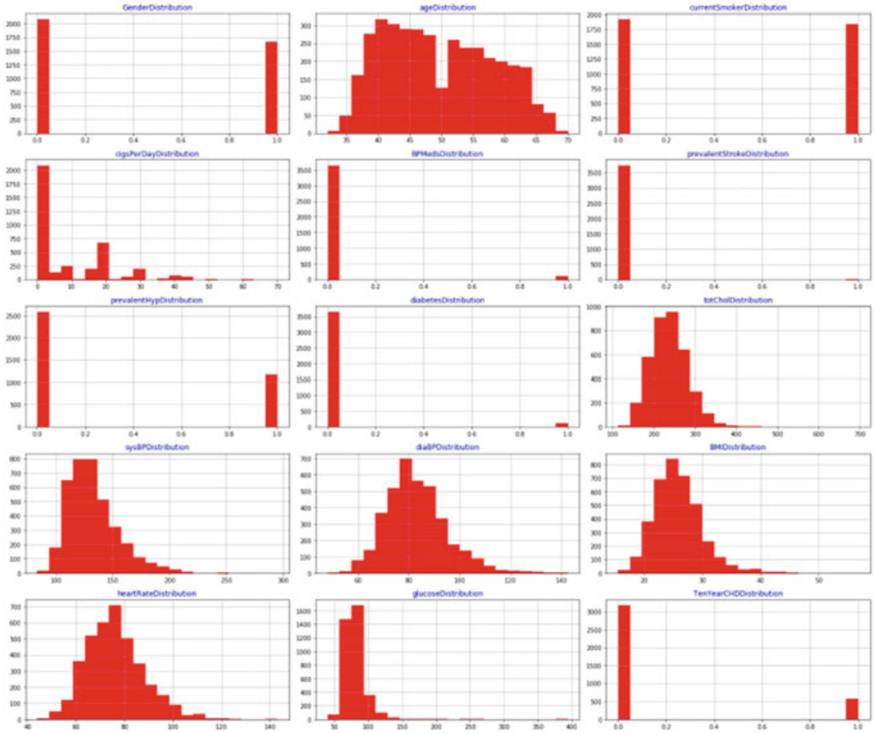


**Fig. 4** Attribute details

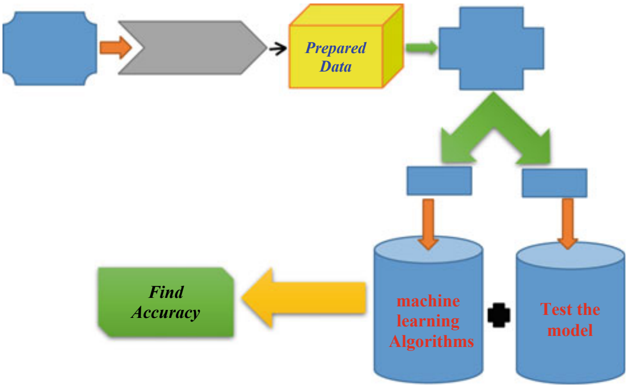


**Fig. 5** Architecture

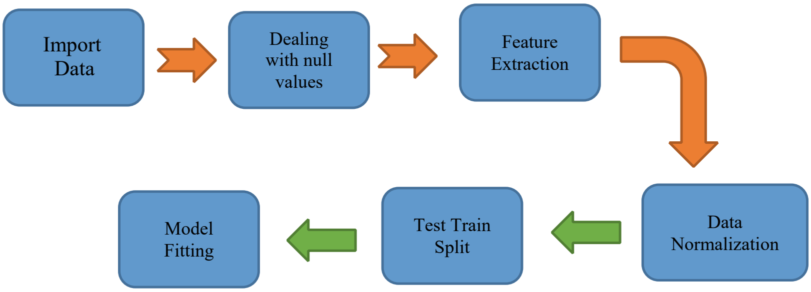


**Fig. 6** Data preprocessing

After normalization and cleaning, the data are split into training and testing data, which are done in the ratio of 7:3, which means that 70% of the data is used for training and 30% of the data is used for testing. Now, the data are ready to be used for classification (Fig. 6).

## 6 Accuracy

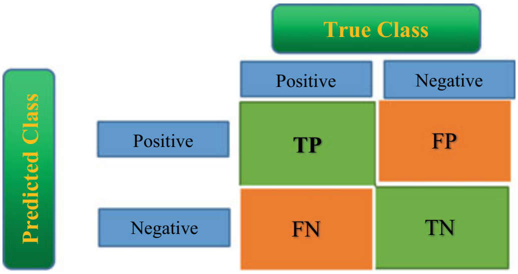
Accuracy is a parameter that describes how correct the model is. It is basically ‘100-error’. A confusion matrix table is used to determine a model’s correctness (Fig. 7).

Here, the terms True Positive (TP) and False Positive (FP) denote actual positive and anticipated positive results, respectively. False Negative (FN) denotes actual positive and predicted negative results, and True Negative (TN) denotes the opposite (Actual negative, predicted negative) [11–20].

The accuracy will be:

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{7}$$

**Fig. 7** Confusion matrix



## 6.1 *Results and Discussion*

The objective of this paper is to find out whether a person is a risk of coronary heart disease or not for the next ten years. This also intends to pinpoint the ratio of patients. It is thus analyzed that 16 out of 100 patients are having a risk of coronary heart disease because of their habits and medical history (Fig. 8). The accuracy of each model is also calculated and compared to find the best out of all the implemented models. The four algorithms are decision tree, logistic regression, random forest, and ANN.

The experiment is performed on medical data of Framingham people which has information about their habits, lifestyles, and medical history. It is thus found that random forest outperforms all other models with an accuracy of 84.89%, whereas logistic regression has an accuracy of 84.81%, decision tree has an accuracy of 84.66%, and ANN has an accuracy of 84.54% (Fig. 9). Random forest performs the best as it is ensemble learning, a method that combines various classifiers to address a challenging issue.

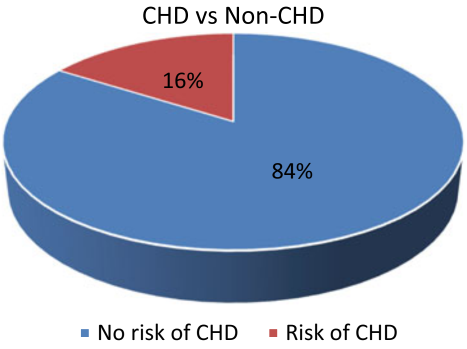


**Fig. 8** Patients' graph

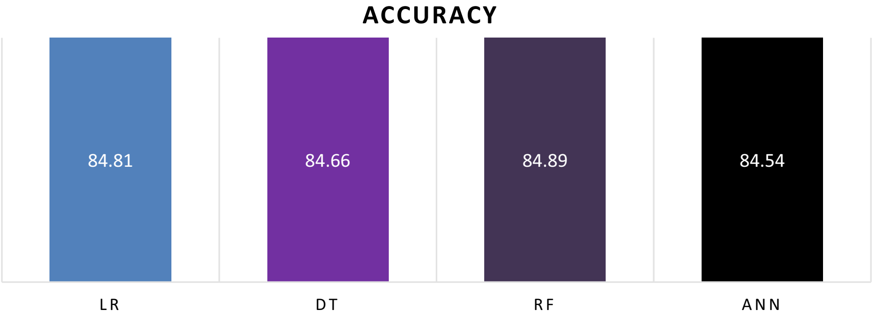


**Fig. 9** Accuracy comparison

## 7 Conclusion and Future Work

In this paper, four supervised classification algorithm of machine learning is used on the data to predict whether a person is having heart disease risk or not. As heart disease is the most rapid cause of death, yet manageable, hence these machine learning algorithms can do a great help in the early and easy prediction, without undergoing painful and risky tests. From this experiment, it is thus found that the random forest is having the highest accuracy out of all the four models. In the future, it is intended to work on advanced evolutionary techniques of an artificial neural network, convolution neural network, and others, to improve the accuracy. Artificial neural with adjusted weights, threshold, and activation function can be a great scope for higher accuracy.

## References

1. Rajdhan A et al (2020) Heart disease prediction using machine learning. Int J Res Technol 9(04):659–662
2. Shah D, Patel S, Bharti SK (2020) Heart disease prediction using machine learning techniques. SN Comput Sci 1(6):1–6
3. Ramalingam VV, Dandapath A, Karthik Raja M (2018) Heart disease prediction using machine learning techniques: a survey. Int J Eng Technol 7(2.8):684–687
4. Mythili T et al (2013) A heart disease prediction model using SVM-decision trees-logistic regression (SDL). Int J Comput Appl 68(16)
5. Katarya R, Srinivas P (2020) Predicting heart disease at early stages using machine learning: A survey. In: 2020 International conference on electronics and sustainable communication systems (ICESC). IEEE
6. Amudhavel J et al (2016) Recursive ant colony optimization routing in a wireless mesh network. In: Proceedings of the second international conference on computer and communication technologies. Springer, New Delhi
7. Kumar MN, Koushik KVS, Deepak K (2018) Prediction of heart diseases using data mining and machine learning algorithms and tools. Int J Sci Res Comput Sci, Eng Inf Technol 3(3):887–898
8. Gavhane A et al (2018) Prediction of heart disease using machine learning. In: 2018 second international conference on electronics, communication and aerospace technology (ICECA). IEEE
9. Bhatla N, Jyoti K (2012) An analysis of heart disease prediction using different data mining techniques. Int J Eng 1(8):1–4
10. Marimuthu M et al (2018) A review on heart disease prediction using machine learning and data analytics approach. Int J Comput Appl 181(18):20–25
11. Chiuchisan I, Geman O (2014) An approach of a decision support and home monitoring system for patients with neurological disorders using internet of things concepts. WSEAS Trans Syst 13(1):460–469
12. Soni J et al (2011) Intelligent and effective heart disease prediction system using weighted associative classifiers. Int J Comput Sci Eng 3(6):2385–2392
13. Yuce MR (2010) Implementation of wireless body area networks for healthcare systems. Sens Actuators A: Phys 162(1):116–129. 13 Heart Disease Detection Using Machine Learning Techniques
14. Ghadge P et al (2015) Intelligent heart attack prediction system using big data. Int J Recent Res Math Comput Sci Inf Technol 2(2):73–77

15. Saw M et al (2020) Estimation of prediction for getting heart disease using logistic regression model of machine learning. In: 2020 International conference on computer communication and informatics (ICCCI). IEEE
16. Bittencourt HR, de Oliveira Moraes DA, Haertel V (2007) A binary decision tree classifier implementing logistic regression as a feature selection and classification method and its comparison with maximum likelihood. In: 2007 IEEE international geoscience and remote sensing symposium. IEEE
17. Demidova L, Ivkina M (2019) Defining the ranges boundaries of the optimal parameters values for the random forest classifier. In: 2019 1st International conference on control systems, mathematical modelling, automation and energy efficiency (SUMMA). IEEE
18. Jin N, Liu D (2008) Wavelet basis function neural networks for sequential learning. IEEE Trans Neural Netw 19(3):523–528
19. Ghosh P et al (2021) Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques. IEEE Access 9:19304–19326
20. Singh M et al (2016) Building a cardiovascular disease predictive model using structural equation model & fuzzy cognitive map. In: 2016 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE
21. Sharma R, Panigrahi A, Garanayak M, Chakravarty S, Paikaray BK, Pattanaik L. Tomato leaf disease detection using machine learning

# Evaluation of Spatiotemporal Fetal Cardiac Imaging Using Deep Learning Techniques

**Dipak Kumar Nidhi, Khushboo Srivastav, Jukka Heikkonen, and Rajeev Kanth**

**Abstract** Fetal echocardiography is a standard diagnostic tool used to evaluate and monitor fetuses with a defective heart associated with a number of fetal conditions. Deep learning is a machine learning technology which can perform specific tasks with specific goals. Deep learning techniques are used to evaluate fetal cardiac ultrasound cine loops and improve the evaluation of fetal abnormalities. Convolutional neural network along with recurrent neural network was applied in this study as CNN + LSTM, CNN + GRU, and 3D CNN for the processing and classification of ultrasonographic cine loops into various classes. The models were able to sort the fetal cardiac cine loops into five standard views with 92.63%, 94.99%, and 82.69% accuracy, respectively. Furthermore, the models were able to diagnose Tricuspid atresia (TA) and Hypoplastic left heart syndrome (HLHS) with an accuracy of 94.61%, 91.99%, and 86.54%, respectively. These deep learning-based algorithms were found to be effective tools for evaluating and monitoring normal and abnormal fetal heart cine loops.

**Keywords** Deep learning · Fetal cardiac cine loops · CHD lesions · CNN · RNN · 3D CNN

---

D. K. Nidhi (✉)
Tribhuvan University, IOE, Pulchowk Campus, Kathmandu, Nepal
e-mail: nidhi73dipak@gmail.com

K. Srivastav
Department of Obstetrics and Gynecology, Tribhuvan University, People's Dental College and Hospital, (General Hospital Block), Kathmandu, Nepal
e-mail: khushboo.sr@gmail.com

J. Heikkonen
Department of Computing, University of Turku, Turku, Finland
e-mail: jukhei@utu.fi

R. Kanth
School of Information Technology, Savonia University of Applied Sciences, Kuopio, Finland
e-mail: rajeev.kanth@savonia.fi

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_25

## 1 Introduction

Congenital heart disease (CHD) represents the wide range of anomalies present since birth, affecting the typical structure of the fetal heart and resulting in its defective functioning [1]. CHD largely contributes to infant morbidity and mortality due to congenital anomalies. It is estimated that there are around 4–13 cases of CHD per thousand live births [1]. According to the World Health Organization's data (WHO), heart anomalies contributed to 42% of infant deaths from 1950 to 1994 [2]. It has also been observed that heart defects affecting the fetal cardiac structure are often missed in routine antenatal fetal scanning [3]. Precise antenatal diagnosis offers potential clinical benefit with regards to later prognosis. This is true especially in cases where some form of intervention or surgery can be provided in utero itself (e.g., in utero aortic valvuloplasty for HLHS). These potential benefits all rely on an accurate fetal diagnosis of CHD. A routine fetal anomaly scan is recommended for all pregnant mothers between 18 and 22 weeks. This includes a basic examination of all the primary organ systems of the fetus to rule out any structural defect [4].

The assessment of fetal echocardiography using cine loops is based on computer-assisted analysis of video frames in order to automatically monitor and evaluate fetuses without the need for manual intervention. Deep learning is a machine learning technology that generates feature maps based on artificial neural networks that can do specialized jobs with specified goals [5]. Deep learning has achieved tremendous progress in the disciplines of computer vision [6]. However, enough uses of these techniques have not been achieved so far since it is a comparatively new process. Meanwhile, due to the diagnostic difficulty posed by a small and fast-beating fetal heart and physician's relative little exposure to congenital heart diseases, fetal echocardiography is very complicated.

This study hypothesized that using input data curated according to clinical guidelines (i.e., selecting only the five screening cardiac views) would allow our models to detect diagnostic signals on small datasets and classify it into the five standard fetal cardiac planes. It would further differentiate structurally normal fetal heart from the two significant congenital heart abnormalities, i.e., Tricuspid atresia (TA) and Hypoplastic left heart syndrome (HLHS) [7].

## 2 Related Works

Hinton introduced deep Belief Networks (DBNs) as a superior method of training each network layer [8]. The ignited deep neural networks research resulted in numerous deep learning-based video classification models. A video clip was seen as a sequence of several frames in specific chronological order. Frame classification, a deep learning approach, is one of the most common video classification methods [9]. In this method, features of each frame were extracted using convolutional neural network (CNN). As input to a classifier, all frame-level parameters were averaged into a video-level representation. Zha [9] investigated the effectiveness of image-based video pattern identification using convolutional kernels for classification and features from various deep learning model layers. For adequate performance, they demonstrated this by fusing CNN features with kernel SVMs. Frame classification methods do not account for the relationship between frames and time, which results in loss of temporal aspect and motion information.

CNNs are very successful in a wide range of activities. As a result, in contrast to image-based classification techniques, several works specialize in using CNN models to discover hidden spatiotemporal feature patterns in extremely long videos. The temporal features in a video were used by end-to-end CNN architectures. Ji [10] extended the conventional 2D CNN meant for pictures to the spatiotemporal domain by creating a 3D CNN model that works on stacked video frames. In order to improve upon the conventional 2D CNN architecture for images, Li [11] created a 3D CNN algorithm that operates on a stacked series of video frames. In 3D CNN, 3D kernels are used to extract spatiotemporal features. CNN training with 3D volume inputs is typically time-consuming. Ji [10] proposed the factorized spatiotemporal CNN approach, which factors the initial learning of the 3D convolution kernels into a sequential process of learning the 2D spatial kernels in the bottom layer to effectively handle 3D input.

Yi [12] described a two-stream temporal CNN that collected motion information between consecutive frames. However, the movements were only shown for a brief period of time. Because complex events in movies generally consist of numerous activities taking place over a lengthy period of time, the above-mentioned approach was insufficient for video analysis. This has motivated researchers to use RNN models in order to account for the sequential or temporal changes in movies. LSTM is one such RNN-based model that provides a good fit while avoiding the “vanishing gradient” effect [13]. Its performance has been proven in a number of tasks, including captioning for images and videos. S. Dash used LSTM model to forecast COVID-19 outbreak [14]. Similarly, S Jain used RNN and CNN in her work to predict soccer result [15]. Akilan developed a combination of LSTM and CNN model outputs to represent spatiotemporal cues for movie classification simultaneously [16]. CNNs and LSTMs are highly complementary, according to the research. Feichtenhofer improved on the two-stream approach by investi- gating a more efficient method of combining spatial and temporal streams [5]. The two-stream approach was found to be superior at modeling correlations of spatial and temporal streams in their study [6, 17].

Kong suggested an automated cardiac phase identification technique based on CNN and RNN [18]. Using an LSTM, the researcher was able to design a temporal regression that detects frames from the end-systolic (ES) and end-diastolic (ED) phases of the cardiac cycle in their study. In contrast, Xue and Debus [19, 20] devised deep learning techniques for left ventricular assessment and quantification utilizing CMR cine loops throughout the cardiac cycle. A deep multitask relationship learning network (DMTRL) for estimating left ventricular indices and identifying cardiac phases was reported in the paper by Xue [19]. In this work, the complicated temporal dynamics of the myocardium were described by an LSTM, which was then applied to enforce the spatiotemporal integrity of following frames. The architecture comprises two concurrent LSTMs, the first of which learns left ventricular deformation during the cardiac cycle to estimate heart indices, and the second of which records spatiotemporal changes over subsequent frames to identify heart phases. Debus and Ferrante [20] also carried out the same duties but included a spatiotemporal CNN that extracts spatial and temporal information using 3D convolutional kernels instead of an RNN. These 3D convolutions use the spatiotemporal data from successive frames to increase the accuracy of cardiac phase identification and quantification [19, 20].

## 3 Methodology

### *3.1 Dataset Collection and Description*

Ultrasound (US) is still the diagnostic technique for fetal echocardiography. This method allows for detailed morphological assessment of fetal cardiac structures. Ultrasound uses a physical medium such as air, water, or tissue to propagate high-frequency sound waves from a transducer. Medical ultrasound devices employ pressure pulses with frequencies ranging from 1 to 15 MHz for diagnostic purposes. The examinations are often carried out in real-time and with the ability to observe multiple planes of the fetal heart depending on the location of the transducer.

To determine the health of the fetal cardiac system, this study used ultrasound data from five standard screening images of normal fetal heart and CHDs. With waived consent, all data were de-identified to maintain the privacy of the person.

The data for this study were collected from several ultrasound centers from September 2020 to August 2021. These data were recorded on Samsung HS40, GE Voluson E8, Samsung HERA W10, GE Voluson E10, and GE Voluson S10. It consisted of videos of five different standard views of normal fetal heart and CHD lesions. The 5 canonical views used were—“Abdominal view (ABDO)”,”4 Chamber (4C)”, “3 vessel view (3VV)”, “right ventricular outflow tract (RVOT)”, and “left ventricular outflow tract (LVOT)”. Similarly, congenital heart diseases considered in this study were Hypoplastic left heart syndrome (HLHS) and Tricuspid atresia (TA). There were a total of 564 videos. The videos were captured at a 30 FPS frame rate by fetal medicine experts. The average length of each video was about 6–10 s.

### *3.2 Baseline Models*

**Convolutional Neural Network(CNN)** CNNs are effective at finding spatial features and using them to classify images[21]. They are made up of numerous hidden layers, filters, kernels, or neurons with learnable weights and parameters. Each channel takes a few inputs and convolves them. Max pooling layer and fully connected layers

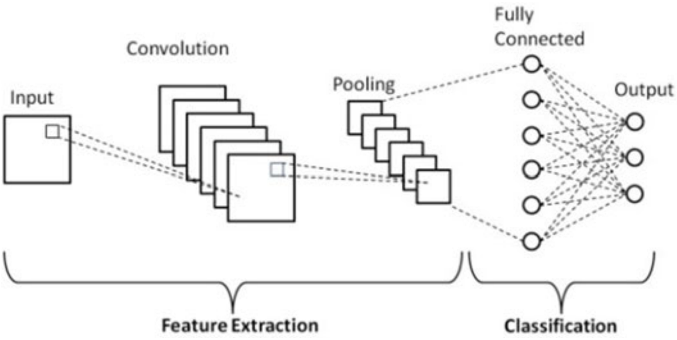


**Fig. 1** Block diagram convolutional neural network (CNN)

come next. Different CNN architectures have been utilized for anatomical structure classification and segmentation in clinical image analysis. Figure 1 displays a block schematic of the CNN architecture.

**Recurrent Neural Network(RNN)** RNNs are a unique class of neural networks that can interpret information and can recognize patterns in ordinal, temporal, or sequential data [6]. However, RNN has a long-term dependency problem and a vanishing gradient problem. To address these issues, LSTM and GRU were used for temporal sequence prediction.

- **Long Short-Term Memory Networks(LSTM)**: LSTM is a form of RNN that is especially developed to handle the issue of long-term dependency[15, 13]. The block diagram of LSTM is shown in Fig. 2.
- **Gated recurrent units (GRUs)**: GRU is a relatively new addition to the recurrent neural networks family[22]. It is very similar to LSTM as shown in Fig. 3.

### *3.3 Main Models*

**CNN + RNN Model:** The phases in the spatiotemporal categorization process are as follows:

1. Data collection.
2. Preprocessing.
3. Spatial feature extraction by CNN.
4. Temporal feature extraction by RNN.
5. Classification.
6. Recognition.

In order to accurately capture the essential information across time, spatiotemporal analysis of fetal cardiac imaging sequences involves the extraction of spatial and temporal features. The CNN + RNN (LSTM/GRU) architecture includes CNN

**Fig. 2** Diagram of long short-term memory (LSTM)

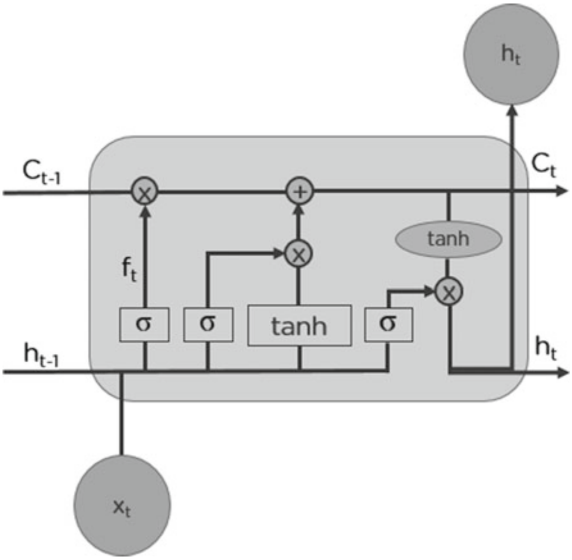
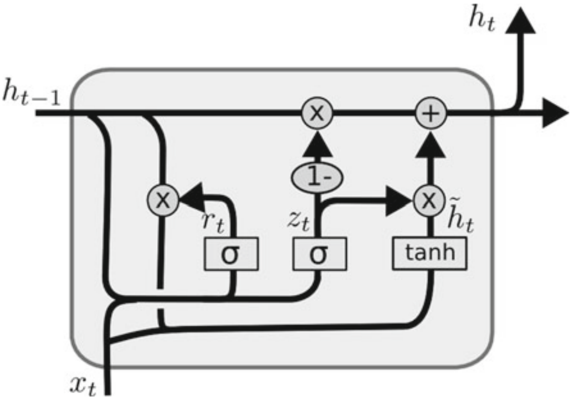


**Fig. 3** Diagram of gated recurrent units (GRUs)



layers for spatial feature extraction on 2D image input data combined with RNN (LSTM/GRU) to support temporal sequence predictions.

**3D CNN Model** 3D CNN is similar to 2D CNN. 2D convolutional neural network (2D CNN) performs well for spatial feature extraction from 2D images, but it does not account for temporal information from the same image. Therefore, 3D CNN is used which is suitable for spatiotemporal feature extraction. Block diagram is shown in Fig. 6. Some of the 3D CNN properties are

1. 3D CNN can model both spatial and temporal features simultaneously.
2. 3D filter of the size $3 \times 3 \times 3$ includes spatial as well as temporal features (Figs. 4 and 5).

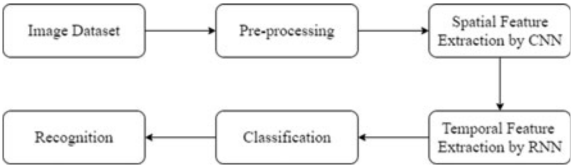


**Fig. 4** System block diagram for CNN + RNN spatiotemporal fetal heart omaging

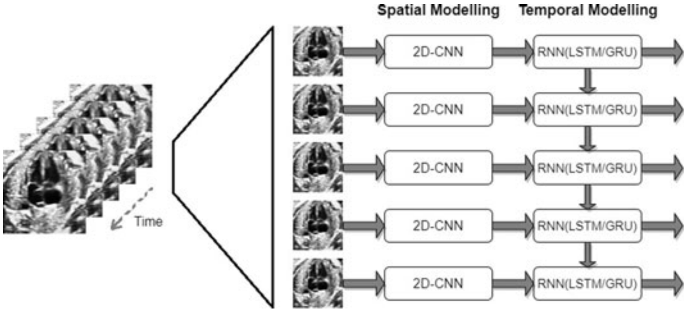


**Fig. 5** Spatiotemporal feature extraction of fetal heart using CNN and RNN (LSTM/GRU)

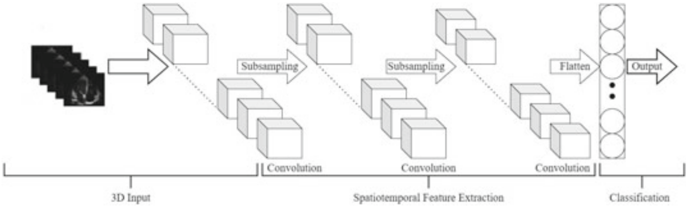


**Fig. 6** Block diagram of 3D CNN

The components of 3D CNN are similar to that of 2D CNN and is given as

- **Convolution Layer**: Voxels are used in 3D CNN in place of pixels which were used in 2D CNN.
- **3D maxpool(2 × 2 × 2)**: In 3D maxpool layer, the maximum value within the volume is selected. In this case, the maximum element within the cube of size 2 is selected.
- **Fully Connected Layer:**: A feed forward neural network is another term for a fully connected layer. The last pooling or convolutional layer's output is flattened and supplied into the fully connected layer as its input.

# 4 Implementation

Figure 7 illustrated the overall system for fetal cardiac cine loops identification and classification along with the prediction of common CHD lesions. 30 FPS cine loops of varying lengths (6–10 s) were converted into images and then passed through the preprocessing pipeline. Image capturing, shuffling, grayscale conversion, resizing, histogram equalization, and rotation were done during the preprocessing method. The preprocessed data was then divided into three groups: “training”, “testing”, and “validation”, in proportions of 60%, 20%, and 20%, respectively. The models CNN + LSTM, CNN + GRU, and 3D CNN were trained using the training datasets, and training loss and training accuracy were calculated after each epoch or training steps. The validation loss and accuracy were also obtained during the training phase. Finally, the performances of the trained models were evaluated with evaluation metrics, confusion matrix, accuracy, sensitivity, specificity, area under the curve AUC), and F1-Score.

## 4.1 Preprocessing

A video is a collection of a set of images called frames, arranged in specific chronological order. For spatiotemporal classification, 30 FPS cine loops of varying lengths (6–10 s) were converted into images then passed through the preprocessing pipeline. During the preprocessing, image capturing, shuffling, grayscale conversion, resizing, histogram equalization, and rotation were done.

**Preprocessing for CNN + RNN Model**: All images collected from different sources, including RGB images were converted to grayscale images (size—150 × 150). The histogram equalization technique was used to improve image contrast by stretching the intensity range. Following that, image augmentation was done. Before resizing the image, the region of interest was extracted to emphasize five standard cardiac planes as shown in Fig. 8.

**Preprocessing for 3D CNN Model**: Videos were of a varying number of frames per second, so it was normalized to a sequence duration of 30 frames. This was

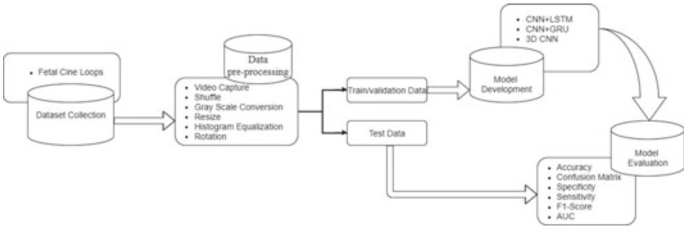


**Fig. 7** Overall proposed system architecture of spatiotemporal fetal cardiac imaging system

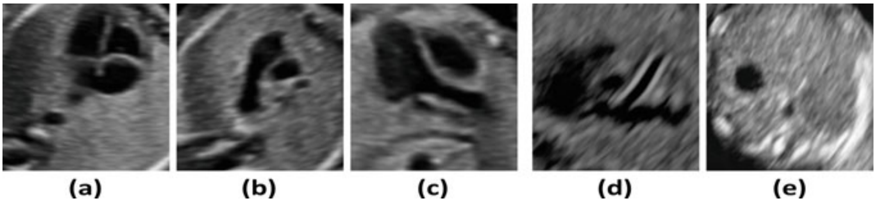


**Fig. 8** Five standard views of fetal heart **a** Four chamber (4 Chamber), **b** three vessel view (3VV), **c** left ventricular outflow tract (LVOT), **d** right ventricular outflow tract (RVOT), **e** abdominal (ABDO)

accomplished by erasing or repeating equidistantly spaced frames. Following that, each video frame was converted into the grayscale frame. Finally, both channels of each sequence (Grayscale and Depth) were normalized to have zero mean and unit variance. This was done so the system could converge faster. Further, images were resized to 150 × 150 × 43 voxel to enable them to fit in the 3D CNN model.

The fetal heart rate fluctuates between 110 and 160 beats per minute. On average, the fetal heart beats 130 times per minute which signifies that one fetal cardiac cycle is completed in 0.47 s. To ensure reliability, 3 cardiac cycles were considered in a sequence averaging 1.41 s. As, there are 30 frames in 1 s, there will be 43 frames in 1.41 s ($1.41 \times 30 = 42.3$). Hence, voxel size of 43 frames was considered in this study.

### *4.2 Development of Models*

The proposed models were developed as 3D CNN, CNN + GRU, and CNN + LSTM for fetal cardiac cine loops identification and classification into five standard planes and the prediction of common CHD lesions.

## 5 Results and Discussion

A fetal cardiac imaging system with fully functional spatiotemporal deep learning models CNN + LSTM, CNN + GRU, and 3D CNN has been developed. The data was divided into three categories: training, testing, and validation, with a 60%, 20%, and 20% split, respectively. The learning rate, optimizer, and number of epochs used for this experiment were 0.001, Adam optimizer, and ten, respectively.

Google Colab and anaconda 4.10.1 have been used to implement the deep learning model. Python programming has been used to create the code. For data preprocessing, opencv Python, NumPy, Pandas, and Theano have been used to generate feature matrix and target vector. The resulting datasets were split into training, validation, and testing using scikit-learn. The Keras and Tensor flow library have been used to implement the spatiotemporal CNN + LSTM and CNN + GRU model in Google Colab. Tensorflow 2.2.0 and keras 2.3.1 have been used to implement 3D CNN model in local anacoda enviroment. Scikit-learn has been used for model evaluation. For graph visualization, the Matplotlib and tensorboard have been used.

### *5.1 Classification and Performance Evaluation of Five Standard View Fetal Heart*

The proposed CNN + RNN(LSTM/GRU) consisted of 7 2D convolutional layers for spatial feature extraction on 2D image input data followed by 2 layers of RNN(LSTM/GRU) to support temporal sequence predication. Similarly, the 3D CNN system consisted of 3 3D convolutional layers for spatiotemporal feature extraction.

The three systems were able to classify 364 fetal cardiac cine loops into five standard fetal heart views, namely ABDO, 4C, 3VV, LVOT, and RVOT.

The overall comparisons were made between implemented 3D CNN, CNN + GRU, and CNN + LSTM models. Table 1 illustrates the performance of the three models. The corresponding bar diagram is shown in Fig. 9.

**Table 1** Comparison of 3D CNN, CNN + GRU, and CNN + LSTM models

| Labels | Training accuracy | Test accuracy | Precision | Recall | Specificity | F1 score | AUC |
|---|---|---|---|---|---|---|---|
| CNN + LSTM | 97.39 | 92.63 | 91.662 | 92.018 | 98.192 | 91.726 | 95 |
| CNN + GRU | 98.02 | 94.99 | 94.174 | 94.84 | 98.79 | 94.436 | 97 |
| 3D CNN | 91.35 | 82.69 | 90.838 | 73.834 | 98.048 | 79.746 | 89 |

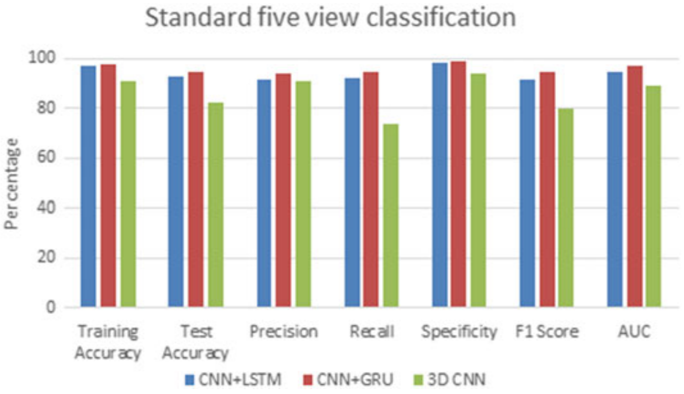


**Fig. 9** Bar chart of performance evaluation of the three models

For five standard views spatiotemporal classification, the CNN + LSTM, CNN + GRU, and 3D CNN models achieved training accuracy of 97.39%, 98.02%, and 91.35%, respectively, and testing accuracy of 92.63%, 94.99%, and 82.69%. CNN + GRU model has higher value of recall (94.174%) and specificity (98.79%). It is known that a high value of recall gives low false negatives and high value of specificity gives high true negatives. Precision and F1-score of CNN + GRU model is also higher compared to the other two models.

The CNN + GRU model performed better than the other two models as its testing accuracy on test datasets was better than the other two models. CNN proved to be good at spatial feature extraction, while RNN (LSTM and GRU) was efficient in temporal feature extraction. The training performance of LSTM and GRU were comparable, but the testing performance of GRU was higher than LSTM. However, GRU had fewer gates than LSTM. Furthermore, 3D CNN's training performance was comparable to the other two models. However, testing accuracy of 3D CNN's model was significantly lower than the other two models.

Hence, CNN + GRU performed better than the other two spatiotemporal models.

### *5.2 Classification and Performance Evaluation of Normal Fetal Heart Versus Congenital Heart Disease (CHD) Lesions*

In this experimental setup, CNN + RNN(LSTM/GRU) consisted of nine 2D convolutional layers for spatial feature detection on the given 2D image data followed by two layers of RNN(LSTM/GRU) for temporal feature extractions. Similarly, the 3D CNN system consisted of three 3D convolutional layers for spatiotemporal feature extraction.

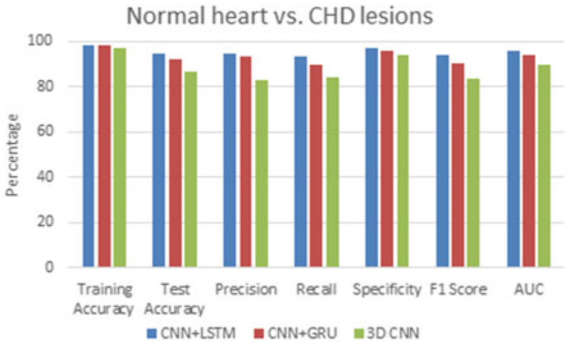
The implemented systems were able to classify 354 fetal cardiac cine loops into five standard fetal heart views, namely three categories, namely normal-4C, HLHS, and TA.

The outcomes of the 3D CNN, CNN + LSTM, and CNN + GRU models for classification of normal fetal heart versus CHD lesions are reported in Table 2. Figure 10 depicts its bar diagram.

**Table 2** Comparison of 3D CNN, CNN + GRU, and CNN + LSTM models

| Models | Training accuracy | Test accuracy | Precision | Recall | Specificity | F1 score | AUC |
|---|---|---|---|---|---|---|---|
| CNN + LSTM | 98.6 | 94.61 | 94.72 | 93.456 | 97.193 | 93.986 | 96 |
| CNN + GRU | 98.63 | 91.99 | 93.22 | 89.626 | 95.69 | 90.713 | 94 |
| 3D CNN | 97.03 | 86.54 | 83.01 | 84.536 | 93.886 | 83.533 | 90 |

**Fig. 10** Bar chart of performance evaluation of the three models



For normal versus CHD lesions spatiotemporal classification, the CNN + LSTM, CNN + GRU, and 3D CNN models gained training accuracy of 98.60%, 98.63%, and 97.30% and testing accuracy of 94.61%, 91.99%, and 86.54%, respectively. Table 2 gives that the overall performance of CNN + LSTM model for normal versus CHD lesions was better than CNN + GRU and 3D CNN. CNN proved to be good at spatial feature extraction, while RNN(LSTM and GRU) was efficient in temporal feature extraction. The training performance of LSTM and GRU was comparable, but the testing performance of GRU was lower than LSTM as GRU had fewer gates than LSTM. Furthermore, 3D CNN's training performance was comparable to the other models, on the other hand, its testing accuracy was significantly lower than the other two models.

Hence, CNN + LSTM outperformed the other two spatiotemporal models.

## 6 Conclusion

Hence, the proposed spatiotemporal deep learning models, i.e., CNN + LSTM, CNN + GRU, and 3D CNN were able to classify the fetal cardiac images into the five standard views with an accuracy of 92.63%, 94.99%, and 82.69%, respectively. CNN + GRU outperformed the other two models in terms of accuracy compared to experimental findings.

Similarly, these algorithms were able to identify and diagnose normal heart versus Tricuspid atresia and HLHS with an accuracy of 94.61%, 91.99%, and 86.54%, respectively. When compared to the experimental findings, CNN + LSTM surpassed the other two models.

## References

1. Ferencz C, Rubin JD, Mccarter RJ, Brenner JI, Neill CA, Perry LW, Hepner SI, Downing JW (1985) Congenital heart disease: prevalence at livebirth: the Baltimore-Washington infant study. Am J Epidemiol 121(1):31–36
2. Crane JP, LeFevre ML, Winborn RC, Evans JK, Ewigman BG, Bain RP, Frigoletto FD, McNellis D, Group RS et al (1994) A randomized trial of prenatal ultrasonographic screening: Impact on the detection, management, and outcome of anomalous fetuses. Am J obstet Gynecol 171(2):392–399
3. Abu-Harb M, Hey E, Wren C (1994) Death in infancy from unrecognised con-genital heart disease. Arch Dis Child 71(1):3–7
4. Salomon LJ, Alfirevic Z, Berghella V, Bilardo C, Hernandez-Andrade E, Johnsen S, Kalache K, Leung K-Y, Malinger G, Munoz H et al (2011) Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. Ultra-Sound Obstet Gynecol 37(1):116–126
5. Feichtenhofer C, Fan H, Malik J, He K (2019) Slowfast networks for video recognition. In: Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp 6202–6211
6. Hernandez KAL, Rienmu¨ller T, Baumgartner D, Baumgartner C (2020) Deep learning in spatiotemporal cardiac imaging: a review of methodologies and clinical usability. Comput Biol Med, p 104 200
7. Paladini D, Volpe P (2018) Ultrasound of congenital fetal anomalies: differential diagnosis and prognostic indicators. CRC Press
8. Hinton GE (2009) Deep belief networks. Scholarpedia 4(5):5947
9. Zha S, Luisier F, Andrews W, Srivastava N, Salakhutdinov R (2015) Exploiting image-trained cnn architectures for unconstrained video classification. arXiv preprint arXiv:1503.04144
10. Ji S, Xu W, Yang M, Yu K (2012) 3d convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell 35(1):221–231
11. Ji S, Zhang C, Xu A, Shi Y, Duan Y (2018) 3d convolutional neural networks for crop classification with multi-temporal remote sensing images. Remote Sens 10(1):75
12. Jasechko S, Sharp ZD, Gibson JJ, Birks SJ, Yi Y, Fawcett PJ (2013) Terrestrial water fluxes dominated by transpiration. Nature 496(7445):347–350
13. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural comput 9(8):1735–1780
14. Dash S, Chakravarty S, Mohanty SN, Pattanaik CR, Jain S (2021) A deep learning method to forecast COVID-19 outbreak. N Gener Comput 39(3):515–539
15. Jain S, Tiwari E, Sardar P. Soccer Result prediction using deep learning and neural networks. In: Intelligent data communication technologies and Internet of Things 2021. Springer, Singapore, pp 697–707
16. Akilan T, Wu QJ, Safaei A, Huo J, Yang Y (2019) A 3d cnn-lstm-based image-to-image foreground segmentation. IEEE Trans Intell Transp Syst 21(3):959–971
17. Arnaout R, Curran L, Chinn E, Zhao Y, Moon-Grady A (2018) Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions. arXiv preprint arXiv:1809.06993
18. Kong B, Zhan Y, Shin M, Denny T, Zhang S (2016) Recognizing end-diastole and endsystole frames via deep temporal regression network. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 264–272
19. Xue W, Brahm G, Pandey S, Leung S, Li S (2018) Full left ventricle quantification via deep multitask relationships learning. Med Image Anal 43:54–65
20. Debus A, Ferrante E (2018) Left ventricle quantification through spatio-temporal cnns. In: International workshop on statistical atlases and computational models of the heart. Springer, pp 466–475

21. Li Z, Liu F, Yang W, Peng S, Zhou J, A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Trans Neural Netw Learn Syst. https://doi.org/10.1109/TNNLS.2021.3084827
22. Cho K et al (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. In: Proceedings of the conference on empirical methods in natural language processing (EMNLP). Association for Computational Linguistics, Stroudsburg, PA, USA, pp 1724–1734

# Application of Deep Learning on Skin Cancer Prediction

**Abhishek Dhiman and Naveen Chauhan**

**Abstract** Skin cancer is one of the deadly diseases that occur on the skin, which may lead to damage in skin tissues, paralysis, and can even cause death of an individual. More accurate the process of diagnosis is and the earlier we do it can reduce the adverse effects of this deadly disease. The model proposed in this paper can help to classify skin cancer just by feeding an image to the model. This study presents a high level of research to differentiate skin lesions from CNN. We limit our reviews to those who are prone to skin lesions. Also, in addition this paper focuses on how the proposed models are quite hard to compare and the challenges that occur while designing such a model. Seeing the large increase in number of cases of skin cancer and the money which is spent on diagnosis, it is quite important to detect the disease as timely as possible. Considering the importance of these problems, researchers have worked on developing a variety of diagnosis strategies for skin cancer. This paper introduces a study of ways to detect skin cancer using deep learning. All the previous works published in conferences and journals were studied. The findings of the study are presented with tools, graphs, tables, strategies, and frameworks for a better understanding. The results obtained of the study done include graphical representation and tabular form of different performance evaluation metrics.

**Keywords** Skin cancer · Deep learning · Transfer learning · Convolutional neural network

---

A. Dhiman (✉) · N. Chauhan
Department of Computer Science and Engineering, National Institute of Technology, Hamirpur, India
e-mail: abhishekdhiman2222@gmail.com

N. Chauhan
e-mail: naveen@nith.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_26

## 1 Introduction

In the last few years, skin cancer is one of the most frequently occurring diseases. It makes quite sense that most of the cancers occur on the skin as it occupies large part of our body. It is usually classified into two types of common cancers: melanoma and non-melanoma. Melanoma is dangerous, a rare, deadly form of skin cancer. American cancer statistics states that melanoma has high rate of deaths per year. Melanoma grows in cells called melanocytes. It can affect anywhere the location of the human body. It usually occurs in areas which are shown to the sun's rays, such as hands, lips, and other parts of body. The type of melanoma cancer can only be cured if it is found early; otherwise, the disease can occur in different body parts and lead to the tragic death of patient. Doctors usually use a biopsy method to diagnose skin cancer. This process removes a sample from the skin that is suspected to be ulcer for medical examination to determine person is suffering from cancer. Taking sample from skin takes a lot of time and patient also gets hurt in this process. Nowadays, advancement in technology provides a comfortable, inexpensive, and quick skin diagnosis symptoms of cancer. In our paper, we have proposed a model that can analyse the skin lesions and classify them by implementing deep learning techniques. We have got better results by implementing transfer learning models like EfficientNet. For testing our model, we used publicly available skin cancer MNIST HAM10000 dataset.

## 2 Literature Survey

In the area of medical image processing, convolutional neural networks are being used frequently. Brinker et al. [1] did study in a systematic manner for doing the classification of skin cancer. The work also discussed the challenges one need to face for classification task. There was a study [2] in which there was a comparison of results of CNN with 58 international dermatologists for the purpose of classification of skin cancer. The CNN had a better performance than most of the dermatologists. The conclusion of the study was that even if dermatologists are very experienced, they can be benefitted from the results provided by image classification done by CNN.

In a study [3], usage of CNN demonstrated that the results obtained are comparable to dermatologists. It also predicted that in near future there will be a rise in smartphone subscriptions for affordable access to necessary diagnostic care. For classifying skin diseases, there has been a lot of use of machine learning techniques. Although not sufficient for advanced classification, retrieval methods are used to complete the basic classification and to evaluate the efficiency of feature engineering. Support vector machines (SVMs), The Bayesian categories, as well as the categories of decision trees, have all been studied. Neural impaired networks use a number of methods that include a multi-horizontal perceptron and back distribution. Convolutional neural networks are an example of their flexibility and predictive power when used in the diagnosis of skin cancer.

In recent years, many applications for classification have been used by CNN. CNNs need large datasets to make accurate predictions. To be used for skin cancer diagnosis, an in-depth study of image classification networks such as Inception, ResNet, AlexNet, VGG was done. It is difficult to compare deep learning structures that deal with the problem of classification of skin cancer, because subjects often use inaccessible or less readily available datasets to be trained and evaluated. Internal datasets are rarely even and often not very similar. Apart from this, the structure of the models is very different, and the test methods may not be compatible.

In [4], authors proposed a new method for handling imbalanced datasets which are highly prevalent in medical datasets. The comparison of state-of-the-art performance was being done on four medical datasets comprising of images which were highly imbalanced. In [5], authors reviewed and compared all the well-known techniques by applying them on publicly available medical datasets. The paper showed an in-depth analysis of performance metrics. In [6], authors proposed a model for challenge created at 2017 International Symposium on Biomedical Imaging. The accuracy obtained was higher than state-of-the-art methods on the same challenge. Also unlike the previous works, value of sensitivity and specificity was also higher which shows the robustness of the model proposed. In [7], authors analysed the relation between accuracy and number of classes as well as the size of dataset. It was observed that the accuracy increases when the size of dataset decreases and the number of classes to predict decreases.

In [8], a study was done in order to compare different deep learning architectures. Techniques like data preprocessing and data augmentation were applied and their experiments achieved accuracy of 92%, recall of 92%, and precision of 93%. In a study done in [9], machine learning approach was proposed in order to classification of skin lesion images. Comparison was made on machine learning classification algorithms like KNN, SVM, and random forest. Results found that SVM was better in comparison with other two.

Most of the research which has been done in this area focuses on binary classification in which we find whether a person is suffering from melanoma or not and not a lot of work has been done for the classification of general images.

## 3 Dataset

The dataset [10] used was publicly available HAM10000 dataset which comprises of total of 10015 images of different types of skin diseases. The dataset comprises of seven different categories of diseases. It comprises of seven diseases whose names are : Melanocytic nevi (6705 images), Melanoma (1113 images), Benign keratosis (1099 images), BCC (514 images), AK (327 images), Vascular Lesions (142 images), and Dermatofibroma (115 images). As we can clearly see that the distribution of these classes is not roughly same, so we applied data augmentation before training the model and learning the features for predicting the skin disease. The image selection done was random.

The size of dataset is a big problem in medical domains as neural networks require data samples whose results we know beforehand for model training. Labelling of medical dataset takes a lot of time and requires expertise in the domain. In other fields, any person can easily perform the labelling of data but the medical field is quite different in that aspect.

## 4 Data Preprocessing

It refers to the period just before the training of model is done. During this period, the input images were resized in order to smoothen the process of training of the model and enhance the classification accuracy. The rescaling of images is done so that to obtain resolution of 64 * 64. This was done in order to reduce the cost of computation occured during processing (Fig. 1).

## 5 Proposed Methodology

### 5.1 Data Augmentation

Data augmentation techniques were made use of in order to solve problems of lesser number of samples of data in the training set, unbalanced distribution of the classes. After application of data augmentation, we obtained balanced sample sizes for each

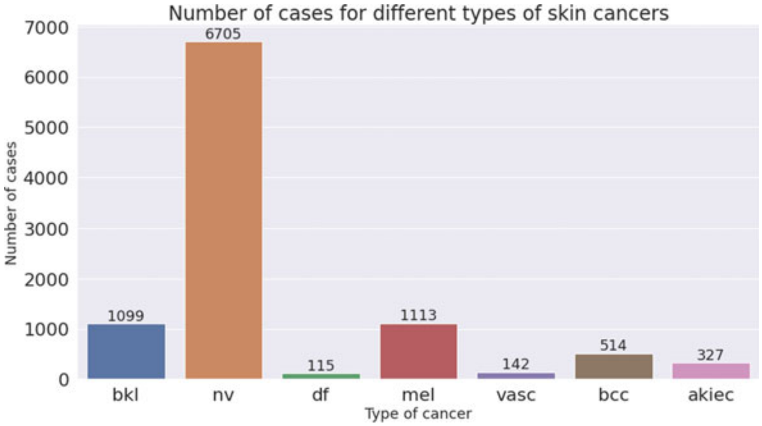


**Fig. 1** Number of instances for different types of skin cancers

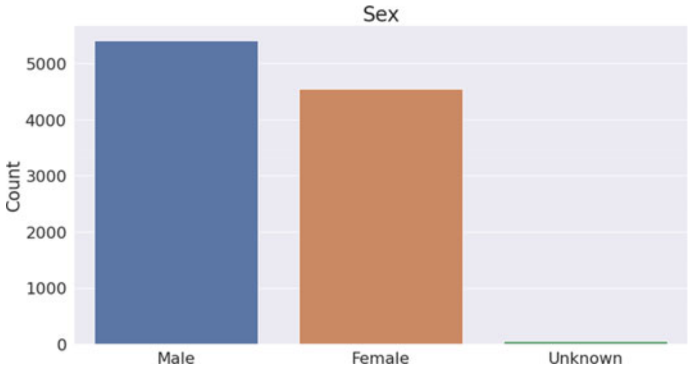


Fig. 2 Number of instances of diseases based on gender

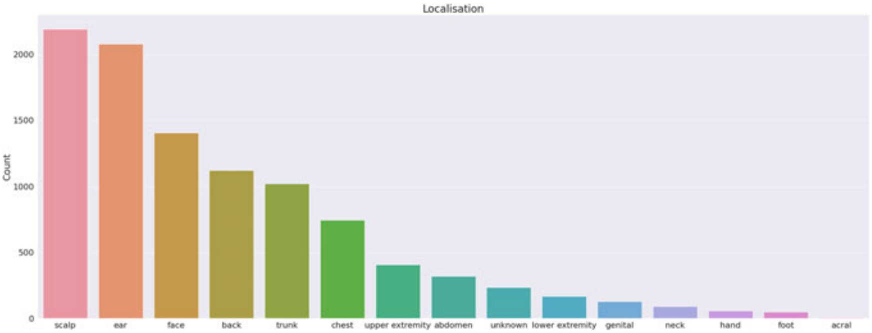


Fig. 3 Number of instances of diseases based on body area

class. In order to achieve data augmentation, the images were rotated by some angle, translated, flipped, zoomed in, or zoomed out. All these mentioned operations were performed on training set only. No data augmentation was performed on testing set (Figs. 2, 3).

## 5.2 Transfer Learning

Transfer learning is an idea in which we use the knowledge gained from one task and extend it to use on other tasks. Observations and parameters learned from training the model on ImageNet dataset are used for solving the problems. Since our model is applied on the medical datasets which are totally different field, we just cannot make use of weights which are trained earlier again to solve the problem and then expect quite good results. So, in order to cope with this we did fine-tuning process in which parameters of model are being changed in a way so that we can apply it to the medical dataset.

### 5.3 Efficient Net Model

Recently, deep learning is being widely used for the purpose of diagnosis of several diseases. Also, lots of research is going on for the classification of several types of skin diseases. Deep learning comprises of several connected layers which use different activation functions and weights. Purpose of activation functions is to adjust the weights. Extraction of features is done by convolutional layers. These layers are basically used to learn and detect complex features in the images. Automatic diagnosis of cancer diseases using images is a difficult task and that is the reason we tried to make use of transfer learning.

The scaled CNN can help to achieve better accuracy. However, earlier research did not made a thorough investigation on the process of scaling. In previous works, it was a common practice to incorporate one of these three—size, width, and depth [11]. It was found that it is better to have roughly equal dimensions. The method of compound scaling was intuitive because bigger the input image, better it is able to detect patterns.

### 5.4 Implementation

**Optimizer**: For the purpose of implementation, Adam optimizer is used. Adam is basically an optimization algorithm which can be used for the purpose of updating the weights. This algorithm is quite different from the stochastic gradient descent in which same learning rate is used for updating all the weights, and there is no change in learning rate during the process of training. Adam is method of adaptive learning meaning that, for different parameters it calculates their individual learning rates.

**Learning Rate**: It refers to change that needs to be done in the model in response to the error generated when the weights of the model are updated. Choosing the value of learning rate is an important task because if we choose very small value for learning rate, it may take forever to train the model, whereas if we choose too large value, we may overshoot the optimum and thus it may lead to an unstable training process.

### 5.5 Stratified Shuffle Split

Stratified shuffle split refers to combination of stratified k fold validation and shuffle split. Firstly, cross-validation is a technique which is used to evaluate machine learning model on given datasets. In this technique, a parameter denoted by $k$ is used which signifies the number of groups in which the given dataset needs to be divided or split. If value of $k$ is 5, we can say it as fivefold cross-validation.

It is used for evaluating the model results on data which is not seen before. This technique gives a less biased estimate of the model and is more robust than train/test split. Now, in stratified k fold, problem is that the data is shuffled only once when the split was made initially and this drastically reduces the chances of overlap of data samples in training and test dataset. Whereas in stratified shuffle split, shuffling takes place every time before the split is made and it preserves the distribution of the classes.

## 6 Performance Evaluation Metrics

### *6.1 Accuracy*

It is the simplest performance metric. It is defined as the ratio of correctly predicted classes of image to the total number of image samples. This metric is valid only when the distribution of classes is almost balanced; i.e., the number of samples of each classes is same.

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

### *6.2 Precision*

It is the ratio of correctly labelled images of positive class to the total number of images which are predicted as positive class.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

### *6.3 Recall*

It is the measure of actual positives predicted by the model. In other terms, it is the ratio of correctly positive labelled to the total number of images inside the positive class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

### *6.4 F1-Score*

Taking look at terms like precision and recall, there seems to be an inverse relationship between the two measurements. If we try to increase precision, recall is decreased and vice versa is also true. Based upon the needs of user and the application for which it is required, we may need to increase one of them. We introduce one more parameter alpha to get more control over the values of precision and recall. This parameter shows how much it prefers one quantity over another. When both the values are having same significance, we set the value of this parameter to one.

To calculate $F1$-score, we need to take harmonic mean of these two values. Higher the value of $F1$-Score, the prediction of model is better. As our problem involves several diseases, making use of $F1$-Score over all the classes provides more better picture of the performance of the model.

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

### *6.5 Confusion Matrix*

Confusion matrix is a 2D matrix which gives a summarized view of predictions performed by the model. It is basically a 2D matrix in which we have the actual label as well as the label which the model has predicted. There are four cases which can occur in confusion matrix: True Positive, False Positive, True Negative, False Negative.

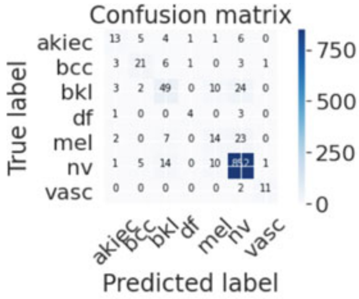
## 7 Results and Discussion

The mean accuracy obtained for five splits was 86.42%. The number of epochs used in each split was 10. In previous works done on the same HAM10000 dataset, P Kanani et al. [12], Nugroho et al. [13], Bassi et al. [14], Moldovan et al. [15], and Cevik et al. [16] achieved an accuracy of 77.98%, 78%, 82.8%, 85%, and 85.62%, respectively.

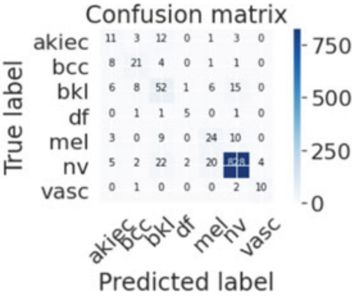
Most of the existing works used train/test split for splitting the dataset. The issue with this is that it can lead to a situation in which some of the class samples are more than others in the split which can result in bias in the model training as it is random split. Thus, to achieve more robust and dependable results, we made use of stratified shuffle split which will preserve the class distribution (Table 1).

**Table 1** Table depicting split-wise weighted average value of precision, recall, and $F1$-score

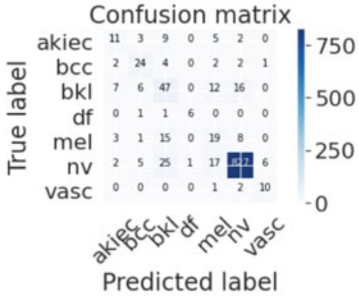
| Split No. | Precision (%) | Recall (%) | $F1$-Score (%) | Accuracy (%) |
|---|---|---|---|---|
| 1 | 86 | 87 | 87 | 87.18 |
| 2 | 87 | 86 | 87 | 86.12 |
| 3 | 87 | 87 | 86 | 85.60 |
| 4 | 89 | 86 | 88 | 86.30 |
| 5 | 88 | 87 | 88 | 86.92 |



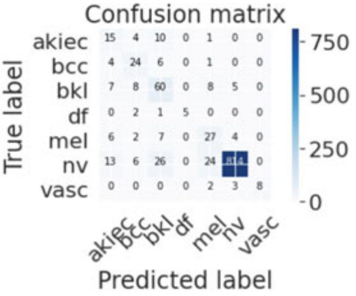
(a) Confusion Matrix for 1st Split



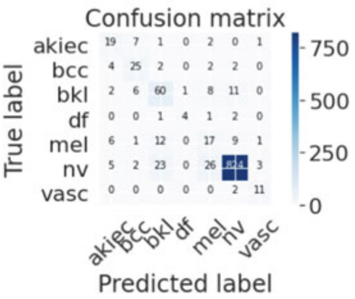
(b) Confusion Matrix for 2nd Split



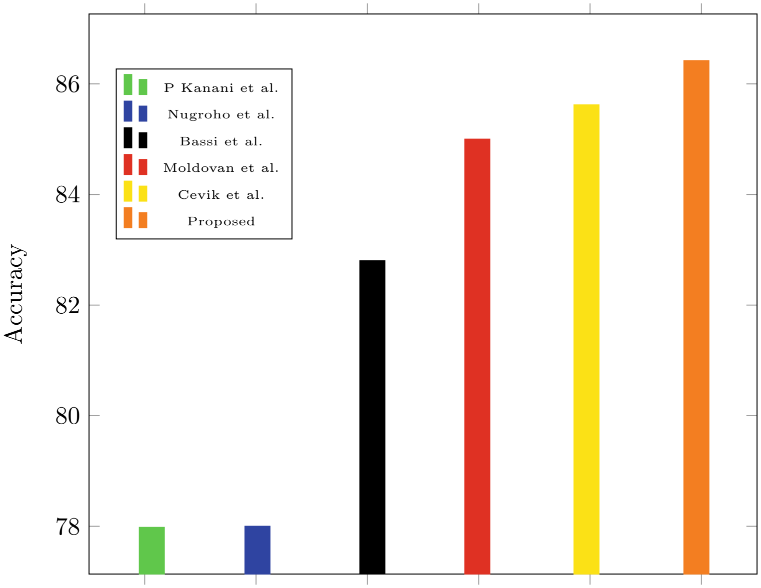
(a) Confusion Matrix for 3rd Split



(b) Confusion Matrix for 4th Split



(a) Confusion Matrix for 5th Split



| Split No. | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| 1 | 86% | 87% | 87% | 87.18% |
| 2 | 87% | 86% | 87% | 86.12% |
| 3 | 87% | 87% | 86% | 85.60% |
| 4 | 89% | 86% | 88% | 86.30% |
| 5 | 88% | 87% | 88% | 86.92% |

## 8 Conclusion

The cases of skin cancer are increasing worldwide each year. Most of the times the skin cancer cases are being checked by dermatologists. But in recent studies, it has been observed that convolutional neural networks are outperforming the dermatologists. So, we applied the efficient net model for the classification of skin cancers on the HAM10000 dataset. We can clearly see that it is very effective in terms of classifying diseases and making use of data augmentation it achieved accuracy of 86.42%. Also, the limitations of the work proposed are that there is a serious class imbalance and in order to cope with this we need more data to have a balanced class distribution for better training of the model.

## References

1. Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, Berking C, Steeb T, Enk AH, Von Kalle C (2018) Skin cancer classification using convolutional neural networks: systematic review. J Med Internet Res 20(10):e11936
2. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Kalloo A, Ben A, Hassen A, Thomas L, Enk A et al (2018) Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 29(8):1836–1842
3. Andre E, Kuprel B, Novoa BA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118
4. Gao L, Lei Z, Chang L, Wu S (2020) Handling imbalanced medical image data: a deep-learning-based one-class classification approach. Artif Intell Med 108:101935
5. Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, Alsaiari SA, Saeed AHM, Alraddadi MO, Mahnashi MH (2021) Skin cancer detection: a review using deep learning techniques. Int J Environ Res Publ Health 18(10):5479
6. Acosta Jojoa MF, Tovar Caballero LY, Garcia-Zapirain MB, Percybrooks WS (2021) Melanoma diagnosis using deep learning techniques on dermatoscopic images. BMC Med Imaging 21(1):1–11
7. Takiddin A, Schneider J, Yang Y, Abd-Alrazaq A, Househ M et al (2021) Artificial intelligence for skin cancer detection: scoping review. J Med Internet Res 23(11):e22934
8. Kassani SH, Kassani PH (2019) A comparative study of deep learning architectures on melanoma detection. Tissue Cell 58:76–83
9. Murugan A, Nair SAH, Kumar KP (2019) Detection of skin cancer using SVM, random forest and KNN classifiers. J Med Syst 43(8):1–9
10. Philipp T, Rosendahl C, Kittler H (2018) The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5(1):1–9
11. Tan M, Le Q (2019) Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR, pp 6105–6114
12. Kanani P, Padole M (2019) Deep learning to detect skin cancer using google colab. Int J Eng Adv Technol Regul Issue 8(6):2176–2183
13. Nugroho AA, Slamet I, Sugiyanto (2019) Skins cancer identification system of haml0000 skin cancer dataset using convolutional neural network. In: AIP conference proceedings, vol 2202. AIP Publishing LLC, p 020039
14. Saksham B, Gomekar A (2019) Deep learning diagnosis of pigmented skin lesions. In: 2019 10th International conference on computing, communication and networking technologies (ICCCNT). IEEE, pp 1–6
15. Moldovan D (2019) Transfer learning based method for two-step skin cancer images classification. In: E-Health and bioengineering conference (EHB). IEEE, pp 1–4
16. Çevik E, Zengin K (2019) Classification of skin lesions in dermatoscopic images with deep convolution network. Avrupa Bilim ve Teknoloji Dergisi 309–318

# Feature Analysis for Detection of Breast Cancer Thermograms Using Dimensionality Reduction Techniques

**Vartika Mishra, Subhendu Rath, and Santanu Kumar Rath**

**Abstract** Breast cancer is observed as the most dominant disease availed in women across the globe. It has been studied that if diagnosed early, the mortality rate can be reduced. Among different modalities such as mammography and ultrasound, thermography is proving to be a successful one which helps to detect the tumor at an early rate. It functions by recording the surface temperature of the required area, and with the help of images consisting of different colors, it helps the doctors to interpret the tumor region. Feature reduction process helps in removing the multicollinearity between the different features thus resulting in less computation time to process. In this study, unsupervised-based nonlinear feature analysis is performed for which different methods are applied based on feature reduction, viz., t-SNE, LLE and KPCA. Different classification models such as random forest and decision tree are applied for classifying between healthy and unhealthy breast. Thus, features obtained by kernel principal component analysis (KPCA) method and classified with random forest give a good accuracy of 89.49%.

**Keywords** Breast cancer · Classification · Feature reduction · Thermography

## 1 Introduction

Among many diseases, cancer is observed to be one such type where the cells in the human body grow abruptly [1]. It can occur in any part of the human body. Among many cancers, one of the most occurring is breast cancer. It is analyzed that it occurs more often in women as compared to men. The cells form a tumor which invades

---

V. Mishra (✉) · S. K. Rath
NIT Rourkela, Rourkela, Odisha 769008, India
e-mail: vartikamishra151@gmail.com

S. K. Rath
e-mail: skrath@nitrkl.ac.in

S. Rath
VCU, Richmond, VA 23219, USA
e-mail: subhendu.rath@vcuhealth.org

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_27

its surrounding tissues [2–4]. Mortality rate can be reduced for breast cancer, if the cancer is diagnosed early. However, many different modalities have been taken into consideration for the detection of breast cancer such as ultrasound, mammography, magnetic resonance imaging (MRI,) X-ray and thermography. These modalities have helped in reducing the mortality rate from 30 to 70% by assisting the radiologists and physicians in diagnosing the abnormalities [5].

Mammography is among one of the modalities which helps in detection of breast tumor. It may help in detecting the cancer for an interval between one and half years to four years [6]. It may increase the future chance of growth in cancer. It does not work well for the dense tissues [7, 8]. Ultrasound is considered as an adjunct to the mammography which is done for women having dense breast tissue. Several research has claimed that ultrasound breast imaging in women with dense breasts and negative mammograms has elated cancer detection rate [9]. Studies have observed that there is lower sensitivity rate for women with dense breast by using mammography. It has been observed that the application of ultrasound screening results in detection of otherwise occult cancer, irrespective of the women having high risk with dense breast or not [9].

Among all the above mentioned modalities, with a decade research, it has been reported that thermography has proved to be one of the early diagnostic modalities. It helps in early detection by 10 years which leads to an immense increment in the survival rate [10]. With the thermal camera, breast thermography detects the tumor by recording the temperature of the surface. By placing the thermal camera in front of the patient, the images are captured appropriately as shown in Fig. 1. This modality is contact-free, non-ionizing, non-invasive and radiation-free in nature [11]. This is suitable for women breast implants, with every age, all breast densities and after surgery monitoring. The vascular heat radiated by that breast surface is measured by the infrared radiation. The increase in local vascularization, the tumor is indicative of high metabolic rate. The normal and abnormal thermogram can be distinguished based on the asymmetric heat pattern between the left and the right breast [12–14]. Physicians analyze the subjects for different abnormalities, but sometimes it becomes difficult to detect the same through naked eyes. So, different computer-aided detection (CAD) systems provide additional assistance in finding the information from the thermograms [15, 16].

Thus, in this study, the work has been carried out on the nonlinear feature reduction techniques for detection of breast cancer between healthy and unhealthy breast which reduces the computational cost as compared to the linear reduction techniques efficiently and hence improving the predictions. Here, while preserving the original distances between the data points, the lower dimensional representation is attained. The transformed images obtained are segmented into left and right breast manually. Further, from these segmented images, gray-level run length matrix (GLRLM) and gray-level co-occurrence matrix (GLCM)-based features are extracted. From the matrices, a total of twenty-seven features are extracted. Further, feature reduction techniques are applied, viz., t-SNE, LLE and KPCA for the features extracted. The different features obtained by applying the reduction techniques are taken for the classification by applying two classifiers, viz., random forest and decision tree.

**Fig. 1** A thermal camera capturing image of the patient and stored as a thermogram

Further, the work is organized as: related work described in Sect. 2. Section 3 describing breast cancer methodology and Sect. 4 discusses experimental result analysis. Section 5 thus concludes the proposed work.

## 2 Literature Survey

For applications of data analytics, large amount of day results in worsening the results; hence, redundant data must be minimized [17, 18]. So, this section emphasizes on the state of the art by different researchers for the detection of breast cancer using thermograms.

Borchatt et al. have detected presence of cancer by using the temperature range on region of interest (ROI), eight-level posterization by quantizing the higher tone, standard deviation and the mean temperature. Further, feature extraction method was applied from the breast quadrants and entire image as well by extracting the aforementioned features. Further, support vector machine (SVM) was applied which obtained 86% accuracy [19]. Milosevic et al. had computed twenty GLCM-based texture features. Further, Naïve Bayes, *k*-nearest neighbor (KNN) and support vector machine (SVM) classifier were applied. The classification performance for breast cancer detection was computed by fivefold cross-validation. Among different classifiers, KNN classifier gave better result as compared to the other two classifiers, viz., SVM and Naïve Bayes [20].

Ali et al. have extracted texture features based on the statistical features first-order and second-order (GLCM) features. They performed and analyzed the accuracy of their model by applying the SVM classifier for normal and abnormal breast. They applied four different scenarios and compared their results. Every scenario consists of different proportions of training and testing images. It was observed that the first scenario performed better than other scenarios. They applied different kernel functions for the different scenarios and obtained a highest accuracy of 85% for statistical features for linear and quadratic kernels. By applying quadratic and polynomial kernel functions for GLCM-based features, they obtained a highest accuracy of 80% [21]. Jakubowska et al. have applied wavelet transform on the breast images. Further, they have applied PCA and LDA for dimensionality reduction methods for reducing the number of features. These transformed images with dimensionality reduction methods are further classified with the multi-layer perceptron classifier [22].

Resmini et al. applied dimensionality reduction technique, viz., principal component analysis (PCA). A total of sixteen features were chosen and further classified with the SVM classifier implementing on WEKA tool. Their methodology attained an accuracy of 82.14% with specificity and sensitivity of 25 and 91.7% [23]. Gogoi et al., applied Mann Whitney Wilcoxon test and further obtained 84.75% of the highest accuracy by applying support vector machine (SVM) [24].

In this work, significant features are extracted from the three feature reduction techniques. The t-SNE technique gives three most important features; LLE gives ten features, and KPCA gives 7 most important features. These obtained set of features are further classified with random forest and decision tree classifiers.

## 3 Breast Cancer Methodology

### *3.1 Dataset*

In this study, the dataset used is from an online source available at Database Mastology Research (DMR), Brazil [25]. The dataset consists of 56 subjects each with 20 different positional temperature matrix among which 19 are healthy and 37 are unhealthy. The Thermal camera FLIR SC-620 is used for capturing the thermal image of the subject.

### *3.2 Proposed Approach*

The breast cancer detection between healthy and unhealthy breast thermograms is carried out as shown in Fig. 2. Primarily, the breast images are obtained from the dataset available in the form of temperature matrix as shown in Fig. 3 in preprocessing. The images are further segmented between left and right breast using the manual segmentation method. Further, by applying GLCM and GLRLM matrix, the features are extracted from the images [26].

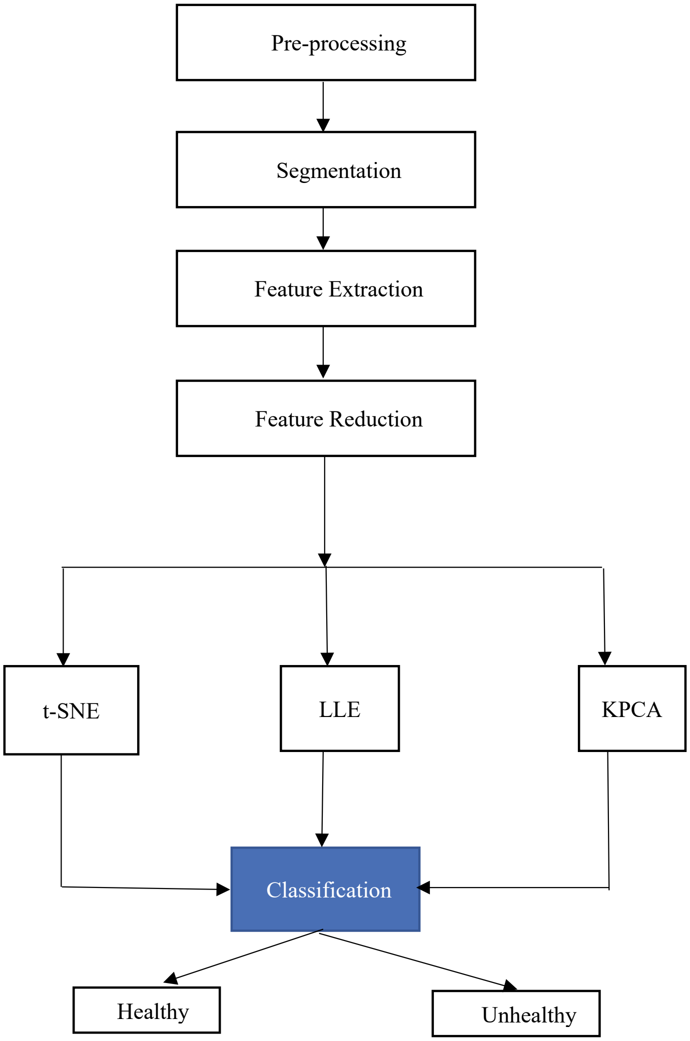


**Fig. 2** Steps of the proposed work for detecting healthy and unhealthy breast thermograms

### 3.2.1 Feature Reduction

Feature extraction often yields the redundant data resulting in worse performance of the models leading to curse of dimensionality. So, this calls for the removal of the features that are less important or irrelevant. This is done by feature reduction techniques.

The resultant of applying dimensionality reduction techniques is a meaningful representation of the reduced data which is obtained by transformation of the high-dimensional data into relevant dimensionality. This reduced dimensionality accounts

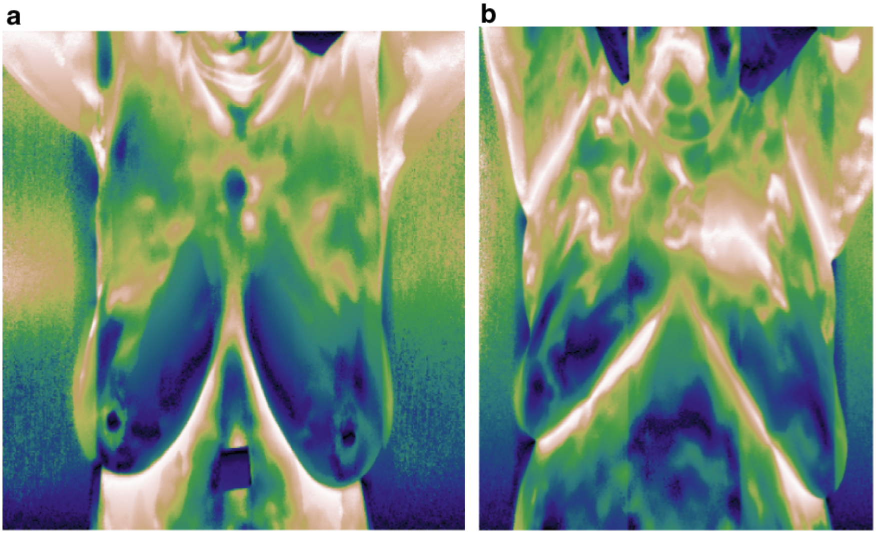


**Fig. 3** **a** A subject with healthy breast thermogram and **b** a subject with an unhealthy breast thermogram

for minimum parameters which are needed for observing the properties of the required data [27]. It facilitates among different fields, viz., visualization, compression of high-dimensional data and classification. This work concentrates on three-dimensionality reduction techniques, viz., t-SNE, LLE and KPCA. Each of them is described as follows:

### a. t Distributed Stochastic Neighbor Embedding (t-SNE)

t-SNE technique simultaneously retains local and global structure of the data. For both, high-dimensional and low-dimensional space, probability similarity points are computed. It is observed that, results obtained from applying t-SNE method are good for linear and nonlinear data [28]. It makes the visualization of the data better and clear. In this work, n components with 3 are taken into consideration for feature reduction.

### b. Locally Linear Embedding (LLE)

LLE starts with finding the nearest neighbors for all sample and further computes set of weights. It is an unsupervised dimensionality reduction method, which preserves the originality of the local features. Further, for finding the lower dimensional embedding points, it applies the eigenvector-based technique [29].

### c. Kernel Principal Component Analysis (KPCA)

It is a nonlinear form of PCA, which helps to solve the more complicated correlation among the different features. The decision boundaries are described by the nonlinear functions. It helps to project the high-dimensional data from the original data with less computational time. This is also known as nonlinear feature mapping [30].

## *3.3 Classification and Performance Parameters*

The classification of breast thermograms is analyzed by applying the tree-based learning algorithms. It is observed that tree-based learning algorithms enhance the models with easy interpretation, stability and high rate of accuracy. It works well for nonlinear relationships and is adaptable for solving problem efficiently.

### 3.3.1 Random Forest

It is a supervised learning algorithm and consists of different trees which leads to ensemble learning method. It measures the importance of feature relatively while making the predictions. The various input variables are recorded enabling the higher number of observations for making predictions [31].

### 3.3.2 Decision Tree

It is a supervised learning method based on conditions for every feature-based split present in the dataset. The decisions are resulted from the tree-like structure. The process starts from the root node and ends with the leaves giving the predictions after computing all the features [31].

### 3.3.3 Performance Parameters

The different parameters of a classification algorithm are visualized by a table known as confusion matrix. It describes the relationship for a set of test data given the true values. The values are determined with the help of four data elements, viz., as described below (Fig. 4).

True Positive (TP): It predicts the true value for the actual true value.

**Fig. 4** A confusion matrix to evaluate the performance metrics

| | | Predict | |
| --- | --- | --- | --- |
| | | **Yes** | **No** |
| Actual | **Yes** | True Positive | False Negative |
| | **No** | False Positive | True Negative |

**Table 1** Different parameters for observing the performance of the random forest classifier

| Classifiers | Sensitivity (%) | Specificity (%) | Precision (%) | $F1$-Score (%) |
|---|---|---|---|---|
| t-SNE | 86.91 | 62.50 | 81.96 | 84.36 |
| KPCA | 89.58 | 89.32 | 94.17 | 91.82 |
| LLE | 91.42 | 61.72 | 84.19 | 87.65 |

**Table 2** Different parameters for observing the performance of the decision tree classifier

| Classifiers | Sensitivity (%) | Specificity (%) | Precision (%) | $F1$-Score (%) |
|---|---|---|---|---|
| t-SNE | 63.83 | 79.77 | 62.76 | 63.29 |
| KPCA | 89.10 | 82.28 | 89.51 | 89.30 |
| LLE | 89.07 | 59.11 | 74.77 | 76.38 |

**Table 3** Accuracy parameter for different set of features obtained from two different classifiers

| Classifiers | Decision Tree (%) | Random Forest (%) |
|---|---|---|
| t-SNE | 74.22 | 78.66 |
| KPCA | 86.57 | 89.49 |
| LLE | 76.38 | 82.22 |

True Negative (TN): It predicts a false value for an actual true value.
False Positive (FP): It predicts a true value for an actual false value.
False Negative (FN): It predicts a false value for an actual false value.

The values of performance parameters, i.e., accuracy, specificity, $F1$-score, sensitivity, specificity and precision are computed based on two different classifier models, i.e., random forest and decision tree as shown in Tables 1, 2 and 3.

## 4 Experimental Result Discussion

In this proposed approach, feature reduction-based work is carried out between three techniques mainly: t-SNE, LLE and KPCA. Each technique has reduced the features to different dimensions: three features from t-SNE, seven features from KPCA and ten features from LLE. These features are further applied individually for classifying among the healthy and unhealthy breast. Decision tree and random forest classification methods are applied, for each set of features obtained from the abovementioned methods for classification.

In Table 1, the accuracy of the two different classifiers is shown which gives an observation that KPCA has given better set of features when classified with random forest classifier with an accuracy of 89.49% as compared to other two dimensionality reduction techniques.

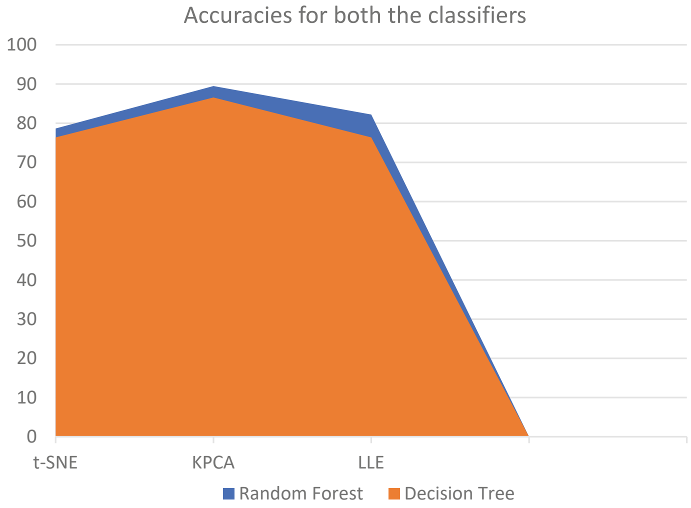


**Fig. 5** A chart representing the accuracies obtained for both the classifiers

Among all the four parameters calculated for the random forest classifier, it is observed that the precision value of the set of features obtained from KPCA reduction technique gives a highest value of 94.17% as mentioned in Table 1. This gives an analysis that the true positive rate calculated from KPCA set of features is good as compared to the other two reduction method. It selects the most significant eigenvectors and eigenvalues giving the low-dimensional representation of the data objects. The sensitivity rate for t-SNE and LLE method is high which determines that the false-positive rate is more. It also accounts for catching good number of actual cases of the disease.

The precision rate for decision tree is high, i.e., 89.51% as compared to other performance parameters for the KPCA set of features obtained. The sensitivity rate for both the other techniques is higher viz. t-SNE and LLE as compared to KPCA set of features as shown in Table 3 (Fig. 5).

## 5 Conclusion

From the above proposed work, it is concluded that kernel principal component analysis feature reduction technique gives better set of features when compared with other two techniques, i.e., t-SNE and LLE with the highest precision rate of 94.17% for decision tree classifier and the highest accuracy of 89.49% for random forest classifier for classifying between healthy and unhealthy breast. In future work, deep learning methods will be applied to obtain better prediction for detection of breast cancer.

## References

1. Cancer statistics, https://www.cancer.gov/about-cancer/understanding/what-is-cancer
2. Breast Cancer, https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html
3. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, CA. Cancer J Clin 67(1):7–30
4. Sree SV, Ng E-K, Acharya RU, Faust O (2011) Breast imaging: a survey. World J Clin Oncol 2(4):171–178
5. Schneider M, Yaffe M (2000) Better detection: improving our chances, digital mammography. In: 5th International workshop on digital mammography
6. Tabár L, Vitak B, Chen H-HT, Yen M-F, Duffy SW, Smith RA (2001) Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer: Interdisc Int J Am Cancer Soc 91(9):1724–1731
7. EtehadTavakol M, Vinod Chandran EYK, Ng, and Raheleh Kafieh, (2013) Breast cancer detection from thermal images using bispectral invariant features. Int J Therm Sci 69:21–36
8. Brown J, Bryan S, Warren R (1996) Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. BMJ 312(7034):809–812
9. Lee CH, David Dershaw D, Kopans D, Evans P, Monsees B, Debra Monticciolo R, Brenner J et al (2010) Breast cancer screening with imaging: recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. J Am Coll Radiol 7(1):18–27
10. Etehadtavakol M, Ng EYK (2013) Breast thermography as a potential non-contact method in the early detection of cancer: a review. J Mech Med Biol 13(2):1330001
11. Tan TZ, Quek C, Ng GS, Ng E-K (2007) A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure. Expert Syst Appl 33(3):652–666
12. Sathish D, Kamath S, Rajagopal KV, Prasad K (2016) Medical imaging techniques and computer aided diagnostic approaches for the detection of breast cancer with an emphasis on thermography—a review. Int J Med Eng Inf 8(3):275–299
13. Lee H, Chen Y-P (2015) Image based computer aided diagnosis system for cancer detection. Expert Syst Appl 42(12):5356–5365
14. Lipari CA, Head JF (1997) Advanced infrared image processing for breast cancer risk assessment. In: Proceedings of the 19th annual international conference of the IEEE engineering in medicine and biology society. Magnificent milestones and emerging opportunities in medical engineering' (Cat. No. 97CH36136), vol 2. IEEE, pp 673–676
15. Arakeri M, Reddy G (2015) Computer-aided diagnosis system for tissue characterization of brain tumor on magnetic resonance images. SIViP 9(2):409–425
16. Moghbel M, Mashohor S (2013) A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. Artif Intell Rev 39(4):305–313
17. Rao A, Noushath S (2010) Subspace methods for face recognition. Comput Sci Rev 4(1):1–17
18. Tomakova R, Komkov V, Emelianov E, Tomakov M (2017) The use of Fourier descriptors for the classification and analysis of peripheral blood smears image. Appl Math 8(11):1563
19. Borchartt TB, Resmini R, Conci A, Martins A, Silva AC, Diniz EM, Paiva A, Lima RCF (2011) Thermal feature analysis to aid on breast disease diagnosis. In: Proceedings of 21st Brazilian Congress of Mechanical Engineering—COBEM, pp 24–28
20. Milosevic M, Jankovic D, Peulic A (2014) Thermography based breast cancer detection using texture features and minimum variance quantization. EXCLI J 13:1204
21. Ali MAS, Sayed GI, Gaber T, Hassanien AE, Snasel V, Silva LF (2015) Detection of breast abnormalities of thermograms based on a new segmentation method. In: 2015 Federated conference on computer science and information systems (FedCSIS). IEEE, pp 255–261
22. Jakubowska T, Wiecek B, Wysocki M, Drews-Peszynski C, Strzelecki M (2004) Classification of breast thermal images using artificial neural networks. In: The 26th annual international conference of the IEEE engineering in medicine and biology society, vol 1. IEEE, pp 1155–1158
23. Resmini R (2011) Análise de imagens térmicas da mama usando descritores de textura. Computer Institute, Master's Thesis, Federal Fluminense University

24. Gogoi UR, Bhowmik MK, Ghosh AK, Bhattacharjee D, Majumdar G (2017) Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In: 2017 International conference on innovations in electronics, signal processing and communication (IESC). IEEE, pp 39–44
25. Dataset Retrieved from http://visual.ic.uff.br/en/proeng/thiagoelias/
26. Mishra V, Rath SK (2021) Detection of breast cancer tumours based on feature reduction and classification of thermograms. Quant InfraRed Thermography J 18(5):300–313
27. Fukunaga K (1990) Introduction to statistical pattern recognition. Academic Press Professional Inc., San Diego
28. Devassy BM, George S (2020) Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE. Forensic Sci Int 311:110194
29. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500):2323–2326
30. Schölkopf B, Smola A, Müller K-R (1998) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10(5):1299–1319
31. Nguyen C, Wang Y, Nguyen HN (2013) Random Forest classifier combined with feature selection for breast cancer diagnosis and prognostic

# Covid-19 Question-Answering System Based on Semantic Similarity

**Shrutikirti Singh and Seba Susan**

**Abstract** Virtual assistants are now a part of our daily lives and have replaced humans in most web-based query systems. The challenge in designing such conversational agents is to replace scripted responses with answers that mimic the way a human communicates with another human. In this work, the task is to create an automated virtual Covid-19 QA agent that would answer user queries related to Covid-19. Since the outset of the Covid-19 pandemic and resulting lockdowns, people have started relying on virtual healthcare agents to resolve their health issues. Many a times, questions worded in different ways have the same meaning and a similar response could be generated for all such cases. Computation of a semantic similarity score between sentence pair representations would facilitate improved natural language understanding. The sentence representations are derived from the hidden-state representations of different neural networks. The proposal in this paper is to investigate state-of-the-art LSTM-based frameworks for application to Covid-19 QA systems that involve computation of semantic similarity between sentence pairs. Experiments performed on the benchmark *Covid-19 QA* dataset containing question–answer pairs related to the Covid-19 pandemic prove that the Siamese LSTM model generates the best responses than the baseline models.

**Keywords** Covid-19 chatbot · Question-answering system · Semantic similarity score · Siamese LSTM · BiLSTM with attention · Medical domain knowledge

## 1 Introduction

The design of automated question-answering (QA) agents, such as Cortana, Alexa, and Siri, is trending research nowadays. Automated virtual assistants (VAs) form the human–machine interface for most web-based services. The most common drawback of web-based query systems is that the answers are mostly scripted and cater to a select set of queries put up by the user [1]. Sometimes, the user wants precise

---

S. Singh · S. Susan (✉)
Department of Information Technology, Delhi Technological University, Delhi, India
e-mail: seba_406@yahoo.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_28

answers to a question that may not match any of the scripted responses. The design of QA agents is composed of three stages: processing the query, processing the documents in the corpus, and extracting the answer to the query. The long short-term memory (LSTM) neural network [2] with some upgradation like the attention mechanism [3, 4] has recently marked success in text-generation tasks such as neural machine translation [3], image captioning [5], and factoid question answering [6]. These problems are solved by learning individual sentence pair representations and training over a similarity measure [7, 8]. Advancements in this field include the substitution of LSTMs by transformers in sequence learning tasks [9] and multilingual chatbots [10] to cater to a multi-ethnic society. The proposal in this paper is to investigate three popular neural network-based frameworks for constructing a Covid-19 QA system. The methodology involves the computation of semantic similarity between sentence pairs, which in our case is computed between the user query and the stored questions in the database. The rest of the paper is organized as follows: Sect. 2 describes related research, while Sect. 3 presents the methodology for our experiments. Section 4 analyzes the results obtained, and Sect. 5 concludes the paper.

## 2 Related Work

In order to design an efficient QA system, it is necessary to comprehend the input query and derive an appropriate answer to the query by referring to the stored question-answers in the database. Factoid question-answering systems return a specific fact such as a place or a person as the answer [6]. These systems constitute of a question categorization step, followed by retrieval of relevant answers from documents or passages. Some traditional strategies of QA systems are pattern matching [11], question similarity matching using simple word-based features [12], Bayesian techniques [13], measurement of answer redundancy on the web [14], and logic-based methods [15].

The long short-term memory (LSTM) neural network model has been used successfully in the past for question-answering tasks [16]. Since 2014, QA systems are popularly implemented using an LSTM as the encoder and another LSTM as the decoder by a technique known as sequence-to-sequence (seq2seq) learning that probabilistically predicts each word in the target response [3]. Variants that have substituted convolutional layers instead of LSTM also exist [4]. Bidirectional LSTM (BiLSTM) processes the inputs in both the forward and the reverse directions, thereby generating two sets of hidden-state vectors [3]. These models perform better than single-direction LSTMs. BiLSTM with attention mechanisms better distinguishes between candidate answers or responses as proved by previous works on sequence learning for dialogue generation [16]. The attention mechanism involves computation of a context vector from the hidden-layer representation extracted from the source sentence which is the user utterance. The context vector highlights the contribution of the neighboring words in the source sentence that would create a better interpretation of the current word encoding. Deep bidirectional RNNs are proposed in [17], where the output of one layer is the input to the second layer. In this paper, we focus on the well-explored LSTM architectures that compute semantic similarity between sentence pairs. The methodology for our experiments is discussed in the next section.

## 3 Methodology

The baseline QA system is the encoder-decoder model [18] popularly implemented using LSTMs. Some form of attention mechanism is usually incorporated [3, 4]. Some popular neural network-based approaches for implementing QA systems are outlined in Table 1.

Since the outset of the Covid-19 pandemic and resulting lockdowns, people have started relying on virtual healthcare agents to resolve their health issues. Many a times, questions worded in different ways have the same meaning and a similar response could be generated for all such cases. Computation of a similarity score would facilitate improved natural language understanding. The sentence representations are derived from the hidden-state representations of different neural networks. The proposal in this paper is to investigate state-of-the-art LSTM-based frameworks for application to Covid-19 QA systems that involve computation of semantic similarity between sentence pairs, which in our case, are the user query and the stored questions in the database.

A few mathematical functions of LSTMs [2] are discussed next. In LSTM, there are three gates: input gate $i$, output gate $o$, and forget gate $f$, and a cell memory $c$; the sigmoid function is denoted by $\sigma$. The time step is denoted by $t$. In the sequence $q^{\text{emb}} = \left\{r^{w1}, \ldots\ldots, r^{wm}\right\}$, considering that the hidden-state vector is denoted by $h_t$

**Table 1** Question-answering models using different neural networks

| Author and year | Method | Highlight |
|---|---|---|
| Bao et al. (2020) [19] | HHH: a medical QA system based on BiLSTM with attention and knowledge graphs | Computation of (attention-based) semantic similarity score between query and stored questions in the database |
| Yao et al. (2018) [8] | Convolutional neural network (CNN) | Semantic similarity computation between query and stored questions in the database |
| Mueller and Thyagarajan (2016) [7] | Siamese LSTM | Semantic similarity computation between query and stored questions in the database |
| Santos et al. (2016) [20] | Attentive pooling network | CNN with attention pooling layer |
| Stroh et al. (2016) [21] | Question-answering using deep learning | Deep learning using end-to-end memory networks |

at time step $t$, the following equations of LSTM are derived.

$$i_t = \sigma\left(W_i r^{wt} + U_i h(t-1) + b_i\right) \tag{1}$$

$$f_t = \sigma\left(W_f r^{wt} + U_f h(t-1) + b_f\right) \tag{2}$$

$$o_t = \sigma\left(W_o r^{wt} + U_o h(t-1) + b_o\right) \tag{3}$$

$$c_t = i_t \times \tan h\left(W r^{wt} + U h(t-1)\right) + f_t \times c_{t-1} \tag{4}$$

$$h_t = o_t \times \tan h(c_t) \tag{5}$$

where the symbol × denotes element-wise multiplication, $b_i, b_f, b_o$ are the bias values, $W_i, W_f, W_o$ are the weight matrices associated with the inputs, and $U_i, U_f, U_o$ are the weight matrices associated with the hidden states. Let $E$ be the embedding matrix that vectorizes each word sentence $w_{it}$ of a sentence. Here, the time step $t \in [1, T]$.

$$x_{it} = E w_{it} \tag{6}$$

The hidden-state representations are computed in the forward and reverse directions for BiLSTM as shown in Eqs. (7) and (8).

$$h_{it} = \text{LSTM}(x_{it}) \tag{7}$$

$$h_{ti} = \text{LSTM}(x_{ti}) \tag{8}$$

The tan$h$ function is used for the nonlinear transformation of the hidden-state vectors; this is followed by a normalizing step to yield the word attention $a_{it}$, as shown in Eqs. (9) and (10).

$$u_{it} = \tan h(U_w h_{it} + b_w) \tag{9}$$

$$a_{it} = \frac{\exp(u_{it})}{\sum_t \exp(u_{it})} \tag{10}$$

A semantic similarity score for the BiLSTM with attention architecture was proposed by Bao et al. in 2020 [19] to measure the closeness between two sentences (in our case, the input query and the stored queries in the database). The semantic similarity between two sentences $p$ and $q$ is computed using the attention weight in (10) as

$$\mathrm{sim}(p,q) = e^{-\left\| \sum_i \sum_j a_{ij}^{(p)} h_{ij}^{(p)} - \sum_i \sum_j a_{ij}^{(q)} h_{ij}^{(q)} \right\|} \tag{11}$$

Finding the top-matching questions in the database will help to generate an appropriate answer to the user query. Likewise, in the Siamese LSTM architecture in [7], a semantic similarity score is computed between the individual LSTM representations of the two sentence pairs. In [8], a three-dimensional tensor is constructed out of word embedding that is given as input to a CNN. The CNN produces the sentence representation which is applied for the computation of semantic similarity score with other sentence representations produced by the CNN. In our work, we compare the utility of the three semantic similarity-based QA systems: convolutional neural network (CNN) [8], BiLSTM with attention [19], and Siamese LSTM [7] for constructing a question-answering system for Covid-19.

## 4 Results

In our experiments, we use three different neural network-based models: convolutional neural network (CNN) [8], BiLSTM with attention [19], and Siamese LSTM [7] to implement a medical QA system trained on Covid-19 questions and answers from a benchmark dataset. All three neural networks employ some form of semantic similarity computation in their architecture framework as observed from Table 1 and discussed in Sect. 3. All three models: convolutional neural network (CNN) [8], BiLSTM with attention [19], and Siamese LSTM [7] were trained on the *Covid-19 QA* dataset [22] that is available online at [23]. The text preprocessing is conducted as per the guidelines in [19]. The *train*:*test* split ratio was 80:20, and threefold cross-validation was used. All experiments are performed in Python software on an Intel PC clocked at 2.8 GHz. A 100-dimensional word2vec embedding was used, and the number of epochs was set to ten. The accuracies of the answers generated for the test set by different models are shown in Table 2. Among the three neural networks, Siamese LSTM outperformed the others with 82.42% accuracy of the answers generated for the test set.

**Table 2** Performance scores of different models on the *Covid-19 QA* test set

| Method | Accuracy (%) |
|---|---|
| CNN [8] | 73.69 |
| BiLSTM with attention [19] | 77.65 |
| Siamese LSTM [7] | 82.42 |

## 5 Conclusion

Virtual agents are commonly encountered in web-based services. This work presents the implementation of a virtual Covid-19 QA agent using different convolutional and recurrent neural network architectures. All models involve the computation of the semantic similarity score which measures the match between two sentence pairs, which in our case, are the user query and the stored questions in the database. The experiments on the benchmark *Covid-19 QA* dataset prove that the Siamese LSTM model performed best in terms of the accuracy of the answers generated for the test set. Exploring more attention models forms the future scope of the current work.

## References

1. Shah AA, Ravana SD, Hamid S, Ismail MA (2019) Accuracy evaluation of methods and techniques in web-based question answering systems: a survey. Knowl Inf Syst 58(3):611–650
2. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
3. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473
4. Mallick R, Susan S, Agrawal V, Garg R, Rawal P (2021) Context-and sequence-aware convolutional recurrent encoder for neural machine translation. In: Proceedings of the 36th annual ACM symposium on applied computing, pp 853–856
5. Arora K, Raj A, Goel A, Susan S (2022) A hybrid model for combining neural image caption and k-nearest neighbor approach for image captioning. In: Soft computing and signal processing. Springer, Singapore, pp 51–59
6. Wang M (2006) A survey of answer extraction techniques in factoid question answering. Comput Linguist 1(1):1–14
7. Mueller J, Thyagarajan A (2016) Siamese recurrent architectures for learning sentence similarity. In: Proceedings of the AAAI conference on artificial intelligence, vol 30, no 1
8. Yao H, Liu H, Zhang P (2018) A novel sentence similarity model with word embedding based on convolutional neural network. Concurrency Comput: Pract Experience 30(23):e4415
9. Goel R, Susan S, Vashisht S, Dhanda A (2021) Emotion-aware transformer encoder for empathetic dialogue generation. In: 2021 9th International conference on affective computing and intelligent interaction workshops and demos (ACIIW). IEEE, pp 1–6
10. Dhanda A, Goel R, Vashisht S, Susan S (2021) Hindi conversational agents for mental health assistance. Int J Appl Res Inf Technol Comput (IJARITAC) 12(1–3):12–20
11. Soubbotin MM, Soubbotin SM (2002) Use of patterns for detection of likely answer strings: a systematic approach. In: TREC, vol 52, p 90
12. Harabagiu S, Moldovan D, Pasca M, Mihalcea R, Surdeanu M, Bunescu R, Girju R, Rus V, Morarescu P (2000) Falcon: boosting knowledge for answer engines. In: TREC, vol 9, pp 479–488
13. Ramakrishnan G, Jadhav A, Joshi A, Chakrabarti S, Bhattacharyya P (2003) Question answering via Bayesian inference on lexical relations. In: Proceedings of the ACL 2003 workshop on multilingual summarization and question answering, pp 1–10
14. Brill E, Dumais S, Banko M (2002) An analysis of the AskMSR question-answering system. In: Proceedings of the 2002 conference on empirical methods in natural language processing (EMNLP 2002), pp 257–264
15. Pasca MA, Harabagiu SM (2001) High performance question/answering. In: Proceedings of the 24th annual international ACM SIGIR conference on research and development in information retrieval, pp 366–374

16. Goel R, Vashisht S, Dhanda A, Susan S (2021) An empathetic conversational agent with attentional mechanism. In: 2021 International conference on computer communication and informatics (ICCCI). IEEE, pp 1–4
17. Graves A, Mohamed A-R, Hinton G (2013) Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, pp 6645–6649
18. Cho K, Merriënboer BV, Bahdanau D, Bengio Y (2014) On the properties of neural machine translation: encoder-decoder approaches. arXiv preprint arXiv:1409.1259
19. Bao Q, Ni L, Liu J (2020) HHH: an online medical chatbot system based on knowledge graph and hierarchical bi-directional attention. In: Proceedings of the Australasian computer science week multiconference, pp 1–10
20. Santos CD, Tan M, Xiang B, Zhou B (2016) Attentive pooling networks. arXiv preprint arXiv:1602.03609
21. Stroh E, Mathur P (2016) Question answering using deep learning. In: SCPD, pp 1–2
22. Möller T, Reina A, Jayakumar R, Pietsch M (2020) COVID-QA: a question answering dataset for COVID-19. In: Proceedings of the 1st workshop on NLP for COVID-19 at ACL 2020
23. https://github.com/deepset-ai/COVID-QA/. Accessed on 5 Apr 2022

# Impact of EEG Signals on Human Brain Before and After Meditation

**Soumya Samarpita and Rabinarayan Satpathy**

**Abstract** Nowadays, busy life style and increased stress causes numerous mental illnesses such as depression, anger, fatigue, and aggressive behaviour. Stress directly affects a human brain and reduces different brain activity. To find the solution of this global phenomenon of this medical complication, advanced drugs like anti-depressants, anxiolytics are developed which also creates long-term side effects. Meditation can also be an alternative medicine to avoid these medical complications. For centuries, the brain and meditation have been linked. Electroencephalogram (EEG) waves are now used to evaluate electrical impulses which are being generated by the human brain. EEG signals can be collected with the help of an EEG recorder. Different brain wave signals are generated by various physiological activities. The primary brain wave signals are alpha, beta, theta, delta, and gamma, which are produced by various physiological activities. This paper focuses on the impact of EEG signals on the human brain before and after meditation.

**Keywords** Stress · Meditation · Electroencephalogram (EEG) · Brain wave

## 1 Introduction

Stress is a significant issue in today's human society. It is a growing problem that has become an unpreventable part of our daily lives. Nowadays, stresses are universal problems affecting human beings. During daily work, stress reduces human functionality and causes different health problem. Stress effects on human brain like depression, heart diseases (blood pressure, heart attack), stomach, and lungs. In industry and education, due to work load, stress detection is very important. Mental stress primarily affects the human brain. Mental stress is one of the problems which

---

S. Samarpita (✉) · R. Satpathy
Sri Sri University, Cuttack, Odisha, India
e-mail: soumya.s2020-21ds@srisriuniversity.edu.in

R. Satpathy
e-mail: rabinarayan.s@srisriuniversity.edu.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_29

arise due to psychosocial or physical situations [1]. Human brain activity and structure are also influenced by stress. As a result, early stress identification is critical for preventing disease and lowering the risk of clinical brain damage and other health issues. Human tension is an important concept which is gradually gaining attention in study of medicine, psychology, neuroscience, etc.

The human brain is the major "organ of the body" that functions as the nervous system's coordinating centre [2]. In a human body, the brain is the most powerful organ. Electroencephalogram (EEG) can be used to analyse brain wave activity signals. The brain's EEG signal is frequently used to diagnose various types of brain activity.

One of the popular techniques is meditation which uses simple methods of mental and body training. Meditation is a method of training attention and awareness, as well as achieving a cognitively clear, emotionally peaceful, and stable state, by employing techniques such as mindfulness or focusing the mind on a specific object. The benefits of meditations are to reduce stress, depression, and pain. Since ancient times, meditation has been used as a normal procedure in alternative medicine to help people live stress-free lives. Many considered that Yoga is a variety of meditation. Yoga is about directing one's mind and using its capacity to monitor one's body. The technique of meditation entails a deliberate tranquillity of the mind by the dissociation of ideas that focus on someone's own breathing. Meditation has been shown to have a wide range of beneficial advantages for the human brain. In the human body, the outcome of mental stress can be classified in a variety of ways. The effects are shown in Fig. 1.

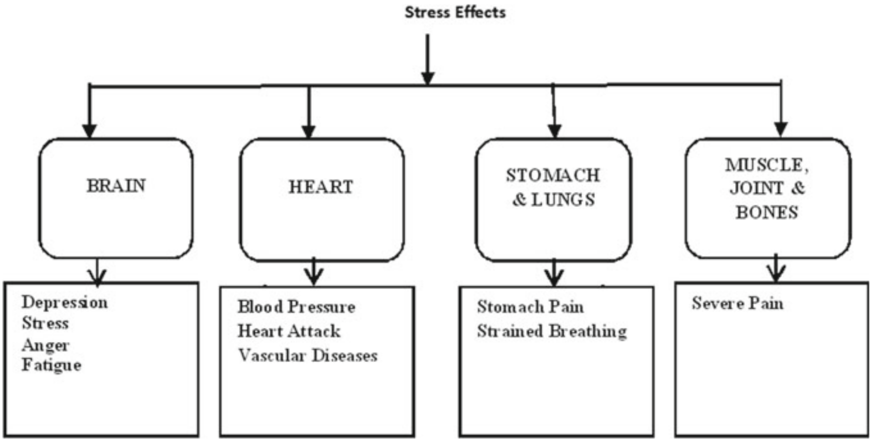


**Fig. 1** General categorization of the outcome of mental stress on the human body [3]

## 2 Connection Between Meditation and Stress

Meditation is the healthy way to reduce stress from human life. Meditation has been shown to be quite helpful in the recovery of those suffering from various addictions. Yoga and meditation are self-regulation techniques used to enhance health, mental disease, stress management, and emotional regulation. In human's body, the most complex organ is the brain. Stress detection uses brain activity as an identifying characteristic. In the study of the human brain, brain–computer interface (BCI) acts as a vital job. Vidal proposed a brain–computer interface (BCI) that bypasses the physiological motor output process and converts brain function that indicates the subject's intention into control signals for a device [4]. The outcome of meditation on human brain is examined using various techniques. Various BCI technologies are used to study the brains which are electroencephalogram (EEG), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). The electroencephalogram signals generated in the brain are acquired and analysed via BCI. Various researchers have been undertaken by exploring the influence of meditation on psychophysiological constants utilising EEG in the current context [2, 5–7]. Monitoring electroencephalogram (EEG) via a brain–computer interface is proposed for experimental investigations into the effects of meditation [8].

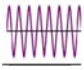
Electroencephalogram which is a standard test that is used to monitor the electrical activity in the human brain can identify signals of brain waves activity [2]. During meditation, oscillations in electroencephalogram (EEG) signal frequencies can be used to map the electrical activity of the human brain. It is also possible to detect potential problems with brain movement. Medical practitioners can use EEG brainwave signals to analyse the brain's activity level, and different meditation techniques can be used to improve mental fitness depending on the health state [6]. Meditation has endless effects on the cognitive processes of the brain. EEG is a signal which is being used to record mind responses in order to obtain the meditation effect on the brain [9].

Electrodes are often put on a patient's scalp in EEG to find out the electrical activity of neurons in the cerebral cortex. EEG does not often record the job of single neurons, but rather identifies the signals produced when groups of neurons are active at the same moment. EEG gives a visual representation of electrical activity in the brain as waves of changing frequency, amplitude, and shape. The most common method for identifying neurological disorders is to use EEG readings [10].

### *2.1 Brain Wave*

Brain wave frequency is estimated in "Hz," i.e. "cycles per second," and it is separated by distinct frequency bands, such as fast waves, moderate frequency waves, and slow waves. As a result, the frequencies of the human brain are separated by distinct bandwidths; various neural activities are described [11].

**Table 1** Classification of brain waves with frequency and state

| Different waves | Frequency range | State | EEG wave samples |
|---|---|---|---|
| Gamma | More than 35 Hz | Gamma brain waves are thought to be involved in conscious attention |  |
| Beta | 13–35 Hz | Beta waves occur when a person is in a state of heightened awareness |  |
| Alpha | 8–13 Hz | Alpha waves occur when a person is temporarily inactive but alert |  |
| Theta | 4–8 Hz | When a repetitious task becomes autonomous, requiring little or no focus to execute, theta waves occur |  |
| Delta | 0.5–4 Hz | Delta waves occur when a person is in deeply sleeping |  |

EEG waves in various mental states are as follows:

a. Gamma waves
b. Beta waves
c. Alpha waves
d. Theta waves
e. Delta waves.

The frequencies of various EEG are illustrated in Table 1.

## 3 EEG Analysis Process Before and After Meditation

The process of interpretation of EEG signals before and after meditation is generally classified into 3 types. Figure 2 depicts the analytical procedure of EEG signals.

(a) Pre-processing: It refers to the removal of noise from the data to get closer to the true neural signal in context of EEG data.
(b) Feature Extraction: Typically, feature extraction is calculated from the recorded signal with the help of a "single electrode" or a "combination of electrodes" [12]. For the analysis of EEG signals, a variety of feature extraction techniques are used. Some of these methods include fast Fourier transform (FFT), eigenvector methods (EMs), wavelet transform (WT), time–frequency distributions (TFDs), auto-regressive method (ARM), and so on.
(c) Classification: By collecting information from EEG data, a classification system can help forecast the approximate aspects of a subject's mental state. It is the

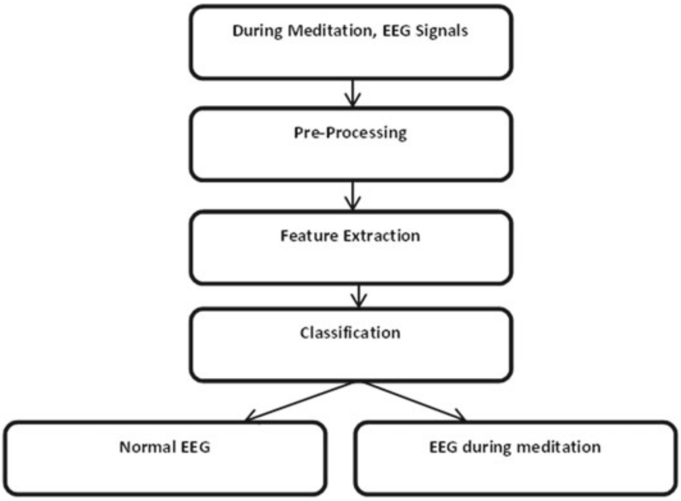


**Fig. 2** Classification of EEG [6]

process of identifying and classifying signals into different classes. Machine learning-based classification has the ability to accurately classify stress states [1]. Some of these techniques include support vector machine (SVM), fuzzy logic, artificial neural network (ANN), and so on.

Signal processing techniques help to advance the field of EEG analysis study. EEG-based investigations often involve recording the physiological signals or brain activity of different subject, extracting frequency domain, time domain, or combined features from these signals and categorising the usable features into several categories [6].

## 4 Comparative Analysis

The evaluation of related literature helps in gaining a better knowledge of the problem and its key features, as well as avoiding duplication. The findings of different studies and experiments encourage the investigator to give importance to useful projects in the field of their studies, and they are able to avoid the past mistakes or defects in the procedure. Having realised the importance of related studies, the investigators will try their best to study the past literature which will be described here in brief.

There is existing research on the effects of meditation on EEG brain waves. This study uses an EEG-based BCI to evaluate the impact of meditation by monitoring brain waves before and after meditation.

In a recent study, an automatic real time of mental stress has been proposed by using frontal lobe EEG. Support vector machine (SVM) and Naïve Bayer's (NB) classifiers are used for detection of mental stress. Amplitude asymmetry, relative power, coherence, absolute power, and phase lag are the key properties retrieved from the recorded EEG signals [3, 13, 14]. Attallah [1] proposed a study for mental state detection using minimum number of fontal brain electrodes. Hybrid feature sets have introduced to detect non-stress and stress state. Principal component analysis (PCA) has been used for feature extraction. Three classifiers such as linear discriminant analysis (LDA), support vector machine (SVM), and K-nearest neighbour (K-NN) have been used to classify the non-stress and stress state. The performance of mental stress detection system is frequently improved by reducing feature space. K-NN and SVM have produced highest accuracies than LDA in their studies.

In [10], discreet wearable gadgets and smart phones were employed to detect stress levels in daily life without annoying the human beings. Alpha wave activities are associated with a peaceful and balanced state of mind, as well as a reduction in stressful situations. HeartMath is a stress relieving smartphone app used in their study. To increase their stress detection accuracy, researchers used multimodal measures.

The influence of concentration (meditation) on emotional reaction is examined using EEG in [12]. Emotion-specific EEG traits include asymmetry of band power which is theta, alpha, and beta as well as Hjorth features. To explore emotional reaction, researchers used the wireless EMOTIV + EPOC collection system. The study uses EEG data collected from "eight brain regions" such as $F3$, $F4$, $T7$, $T8$, $P7$, $P8$, $O1$, and $O2$ while individuals are exposed to various emotions. The recorded signal of a single electrode or a group of electrodes is usually used to calculate features. The asymmetry of the frequency domain band power and the time domain Hjorth pattern were calculated in their study.

Epilepsy is one of the most widespread neurological diseases that affect people over the world. The recording of electroencephalography (EEG) is a non-invasive method for monitoring and diagnosing epilepsy [15]. The approach is intended to be used as a binary classifier. EEG data are used to detect seizures. Researcher proposed convolutional neural networks (CNNs) feature extraction techniques in their study [16–18]. They compared their approach against other classifiers in terms of accuracy, sensitivity, specificity and came up with a score of 98.65% accuracy, 96.29% sensitivity, and 99.25% specificity.

Another investigation shows that during the constant practice of Yoga and Sudarshan Kriya, the EEG signals were recorded in pre- and post-studies [7]. Statistical parameters such as standard deviation, kurtosis, variance, maximum, minimum, and zero crossing have been calculated from each sub-band using discrete wavelet transform (DWT) and db4 wavelet for analysis in their study. They have used Kruskal–Wallis statistical method for validating the obtained parameters. Statistical parameters were used to categorise subjects as meditators or non-meditators using an artificial neural network (ANN). Their method achieved 87.2% classification accuracy in the experiments.

In another study [19], the researchers looked at electrophysiological signals captured during a concentration task to see how regular meditation training affected them. Training classes were held in a "NOVA School of Science and Technology" classroom that provided a calm, airy, and natural light atmosphere between 5:00 and

7:30 p.m. during their research. Band-pass filtering the obtained EEG data between 5 and 30 Hz was done with a 4th-order Butterworth filter. There was a lot of consistency between the self-assessment and the physiological data, they found. In both cases, there are strong signs that performing meditation on a regular basis helps to reduce depression, stress, and anxiety.

The data have been collected using detachable electrodes and connecting leads from the Biopac MP36 recording equipment in [9]. In their research, they applied the discrete wavelet transform. FFT has been used to calculate the power spectrum density. The power spectrum has been used to determine a variety of statistical features. Then, using power spectrum density analysis, the results of the two groups (meditators and non-meditators) have been compared.

There is a comparative analysis of different articles focusing on the impact of EEG signals on the human brain given in Table 2.

## 5 Proposed Methodology

Meditation has significant impacts on brain activity, as seen by alterations in alpha, beta, and theta brain waves linked to improved memory, insight, mind-set, and anxiety reduction. We may be able to design a mechanism in the future that allows us to generate such a vibration that can significantly impact the frequencies in the EEG waves. The prospective future research might also look towards a periodic analysis on the impact of meditation with more subjects, as well as classification using other methods (SVM and K-NN) that include peripheral physiological factors. Meditation training therapies in neurological disorders populations should be the focus of future study.

## 6 Conclusion

Medical practitioners can use EEG brain wave signals to analyse the exercise levels of the brain, and depending on the health situation, different meditation methods can be used to improve mental fitness. From the preceding extensive review, it has been observed that most of the earliest studies focus on the impact of meditation and Yoga on mental stress on human body like brain, heart, stomach, lungs, and bones and so on. Only, few studies have covered on effect of mindfulness meditation on brain wave activities; but, even these studies have focused on only one or two areas like stress, emotion. In the present study, an attempt has been made to provide an impact of EEG signals on human brain before and after meditation. This research focused on the most recent developments in the extraction and identification of human brain activity patterns during meditation. This is a research report; therefore, we clearly ascertain how it can aid in strengthening a person's mental health and making him tension-free after meditation.

**Table 2** Overview of articles focusing into the impact of EEG signals on human brain

| Reference, year | Aims | Type of yoga | Samples | Methods | Conclusion |
|---|---|---|---|---|---|
| [6], 2021 | To investigate the effect of “Yoga” and “meditation” on brain waves | Yoga-Sudarshan Kriya, meditation | Age: 10–17 years | Feature extraction: Frequency and time domain<br>Classification: K-NN | Improve in verbal memory and spatial power after Yoga<br>More research is needed to classify the EEG signal |
| [3], 2021 | To find out the mental stress using frontal lobe EEG signals with the help of automatic real-time detection | Yoga | 14 participants<br>182 samples<br>Age: 18–23 years | Feature extraction: FFT<br>Classification: SVM and Naïve Bayer’s | High accuracy<br>This study’s limitations include gender research and only one lobe |
| [15], 2021 | To examine the deep feature extraction method for “epileptic EEG brain seizure” classification | | 500 distinct EEG files | Feature extraction: CNN<br>Giving a testing EEG dataset to the trained CNN is part of the classification process | The variable extracted feature sizes that can be obtained by changing the density of the layer sizes will be investigated |
| [20], 2021 | To compare the effect of meditation on EEG signal for drug-resistant epilepsy | Meditation | A single patient (42-year-old female) with drug-resistant temporal lobe epilepsy | | No judgement regarding whether meditation is good or harmful should be drawn for single patient |
| [19], 2021 | To see how meditation training affected signals of electrophysiological collected during a concentration task | Meditation (8 weeks) | 35 students | A 4th-order Butterworth filter was used to band-pass filter the acquired EEG data between 5 and 30 Hz | Alpha wave increased |

(continued)

**Table 2** (continued)

| Reference, year | Aims | Type of yoga | Samples | Methods | Conclusion |
|---|---|---|---|---|---|
| [21], 2021 | To find out the effect of meditation using EEG signals to improve cognitive capacities | Meditation (20 min a day for 8 weeks) | Age: 11–16 years | Dry electrodes are used to record EEG data using ENOBIO 32 channel at 500 Hz | Different functional states are improved with meditation |
| [2], 2021 | To study the effect of brain wave during meditation with various physiological activities | Meditation | Age: 20–30 years, 8 volunteers (4 male, 4 female) | The Butterworth filter was utilised with the help of a MATLAB script | Increased delta waves and stress is reduced |
| [4], 2021 | To investigate the anatomical and functional MRI before and after the BCI test | Mental BCI (1 h) | 11 females | Subjects are moved to the EEG chamber after the “pre-MRI” session and equipped with the EEG cap for the BCI experiment | Future research will need to figure out how these early modifications become long-term brain remodelling |
| [1], 2020 | To detect mental stress state using minimum number of frontal brain electrodes | Meditation | Age: 17–26 years<br>66 participants (47 women + 19 men) | Feature Extraction: Principal component analysis<br>Classification: LDA, K-NN, SVM | K-NN and SVM classifiers produced the highest accuracies for both stress detection and stress levels |
| [10], 2019 | To detect stress in daily life scenarios using “smart phones”-”wearable sensors” | | | Median filters wavelet decomposition method used for filtering the noise.<br>Classification: SVM | Calm and balanced state of mind is always indicated by alpha brain wave |
| [22], 2019 | Impact of Yoga on mental health | Hata Yoga | Age: Under 18 years | | These findings highlight potential barriers to practise for men |

(continued)

**Table 2** (continued)

| Reference, year | Aims | Type of yoga | Samples | Methods | Conclusion |
|---|---|---|---|---|---|
| [7], 2019 | To analyse EEG signal-based classification before and after combined "Yoga" and "Sudarshan Kriya" | Yoga, Sudarshan Kriya (one hour daily for 3 months) | 50 males<br>Age: 18–30 years | Feature extraction: Fourier transform, STFT classification: ANN | By expanding the number of subjects in the study, it can be improved. The wavelet transform can be investigated |
| [8], 2019 | To explore the cognitive effect of mindfulness meditation on mental and physical health parameters using EEG | Raja Yoga Meditation | | Neurosky brain wave kit was used with EEG signal for brain wave analysis<br>The pre-amplifier and digitizer pre-process the EEG signal | During Yoga meditation, alpha and theta waves were more predominant |
| [9], 2018 | To analyse the statistical feature of EEG signals for various wavelet | Meditation | 20 regular meditators<br>20 non-meditators | Wavelet transform is used, and then, the Fourier transform is used to get the power spectrum density | Theta wave is increased during meditation |
| [11], 2017 | To study the BCI for removal of stress from human brain | Yoga, Asana, Music Therapy | 150 members<br>Age: 16–60 years | EEG head cap, EEGs are measure electric pulses | Headset is costly<br>Stress is reduced due to Yoga |
| [12], 2017 | To analyse the effect of meditation on emotional response | Meditation (20 min, 8-week meditation) | Age: 22–27 years<br>11 healthy (6 male + 5 female) | Feature Extraction: Hjorth (feature domain)<br>Classification: K-NN<br>EMOTIV, EPOC + wireless EEG device were used as a method of study | Future work will include an analysis of the effect of meditation on a regular basis using more subjects |

(continued)

**Table 2** (continued)

| Reference, year | Aims | Type of yoga | Samples | Methods | Conclusion |
|---|---|---|---|---|---|
| [23], 2017 | To monitor and detect the EEG signals in human brain during depression | Yoga-Sarvangasana and Viparita Karani | Age: 30–50 years | The data were analysed using EEG waves, which were recorded before Yoga in a stressful state | Stress was reduced after meditation |
| [24], 2016 | To examine the statistical features of EEG signal during normal and meditation condition | Meditation-Kriya Yoga | 23 meditators<br>10 non-meditators | Feature Extraction: PCA | Variance, kurtosis, relative band energy were computed |
| [25], 2014 | To compare the BCI accuracy amongst meditation, music training group and non-meditation group | Music meditation (12-week meditation) | Age: 18–22 years<br>40 UG students (Male) | BCI Test | BCI performance of meditation group is improved than music training group |
| [5], 2014 | To analyse the EEG signals during meditation using wavelet transform | Meditation (Vipassan) | 10 normal adults | The EEG signal is analysed by time–frequency analysis–wavelet transform method<br>Feature Extraction: Daubechies8 wavelet function | Alpha wave is more dominant in meditation than non-meditation |

## References

1. Attallah O (2020) An effective mental stress state detection and evaluation system using minimum number of frontal brain electrodes. Diagnostics 10(5):292
2. Nabilah AN, Azami SK, Dafhalla AKY, Adibah MH, Naimah Y (2021) The effect of meditation on brain relaxation incorporating different physiological activities. J Phys: Conf Ser 1962(1):012059. IOP Publishing
3. Alshorman O, Masadeh M, Heyat MBB, Akhtar F, Almahasneh H, Ashraf GM, Alexiou A (2021) Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection. J Integr Neurosci 1–11
4. Nierhaus T, Vidaurre C, Sannelli C, Mueller KR, Villringer A (2021) Immediate brain plasticity after one hour of brain–computer interface (BCI). J Physiol 599(9):2435–2451
5. Fulpatil P, Meshram Y (2014) Analysis of EEG signals with the effect of meditation. Int J Eng Res Technol (IJERT) 3(6)
6. Kora P, Meenakshi K, Swaraja K, Rajani A, Raju MS (2021) EEG based interpretation of human brain activity during yoga and meditation using machine learning: a systematic review. Complement Ther Clin Pract 43:101329
7. Sharma H, Raj R, Juneja M (2019) EEG signal based classification before and after combined Yoga and Sudarshan Kriya. Neurosci Lett 707:134300
8. Mariappan R, Subramanian MR (2019) Experimental investigation of cognitive impact of yoga meditation on physical and mental health parameters using electro encephalogram. In: Soft computing and medical bioinformatics. Springer, Singapore, pp 129–139
9. Gupta N, Sood N, Saini I (2018) Statistical feature based comparison of EEG in meditation for various wavelets. In: 2018 First international conference on secure cyber computing and communication (ICSCCC). IEEE, pp 73–77
10. Can YS, Arnrich B, Ersoy C (2019) Stress detection in daily life scenarios using smart phones and wearable sensors: a survey. J Biomed Inform 92:103139
11. Tiwari A, Tiwari R (2017) Design of a brain computer interface for stress removal using Yoga a smartphone application. In: 2017 International conference on computing, communication and automation (ICCCA). IEEE, pp 992–996
12. Jadhav N, Manthalkar R, Joshi Y (2017) Effect of meditation on emotional response: an EEG-based study. Biomed Signal Process Control 34:101–113
13. Swain A, Satpathy S, Dutta S, Hamad AA (2021) A statistical analysis for COVID-19 as a contract tracing approach and social network communication management, Int J Comp Appl Tech, 66(3–4)
14. Satpathy S, Mangla M, Sharma N et al. (2021) Predicting mortality rate and associated risks in COVID-19 patients. Spat Inf Res. 29:455–464
15. Woodbright M, Verma B, Haidar A (2021) Autonomous deep feature extraction based method for epileptic EEG brain seizure classification. Neurocomputing 444:30–37
16. Mohapatra S, Satpathy S, Mohanty SN (2022) A comparative knowledge base development for cancerous cell detection based on deep learning and fuzzy computer vision approach. Multimed Tools Appl 81:24799–24814
17. Mohapatra S, Satpathy S, Paul D (2021) Data-driven symptom analysis and location prediction model for clinical health data processing and knowledgebase development for COVID-19. In: Nandan Mohanty S, Saxena SK, Satpathy S, Chatterjee JM (eds). Appl Artif Intel COVID-19. Medical Virology: From Pathogenesis to Disease Control. Springer, Singapore
18. Satpathy S, Nandan Mohanty S, Chatterjee JM, Swain A (2021) Comprehensive claims of AI for healthcare applications-coherence towards COVID-19. In: Nandan Mohanty S, Saxen, SK, Satpathy S, Chatterjee JM (eds) Appl Artif Intel COVID-19. Medical Virology: From Pathogenesis to Disease Control. Springer, Singapore
19. Morais P, Quaresma C, Vigário R, Quintão C (2021) Electrophysiological effects of mindfulness meditation in a concentration test. Med Biol Eng Compu 59(4):759–773

20. Delorme A, Grandchamp R, Curot J, Barragan-Jason G, Denuelle M, Sol JC, Valton L (2021) Effect of meditation on intracerebral EEG in a patient with temporal lobe epilepsy: a case report. EXPLORE 17(3):197–202
21. Gupta SS, Manthalkar RR, Gajre SS (2021) Mindfulness intervention for improving cognitive abilities using eeg signal. Biomed Signal Process Control 70:103072
22. Capon H, O'Shea M, McIver S (2019) Yoga and mental health: a synthesis of qualitative findings. Complement Ther Clin Pract 37:122–132
23. Tiwari A, Tiwari R (2017) Monitoring and detection of EEG signals before and after yoga during depression in human brain using MATLAB. In: 2017 International conference on computing methodologies and communication (ICCMC). IEEE, pp 329–334
24. Shaw L, Routray A (2016) Statistical features extraction for multivariate pattern analysis in meditation EEG using PCA. In: 2016 IEEE EMBS international student conference (ISC). IEEE, pp 1–4
25. Tan LF, Dienes Z, Jansari A, Goh SY (2014) Effect of mindfulness meditation on brain–computer interface performance. Conscious Cogn 23:12–21

# COVID-19 Detection and Classification Method Based on Machine Learning and Image Processing

**Inderpal Singh and Priyanka Chawla**

**Abstract** According to WHO, the 2019 new coronavirus outbreak was declared a pandemic in March 2020, and COVID-19 was declared a worldwide epidemic in March 2020. Coronavirus continues to be the most significant cause of death worldwide today. Because the number of cases is increasing every day, emergency clinics have fixed COVID-19 test units. Also, among the various promising research fields, medical images analysis is a field that offers diagnosis facilities and helps in decision-making for multiple diseases, including coronavirus. In the present COVID-19 pandemic, the main challenges include non-COVID-19 cases accurate separation in the early stages of the disease and at the lowest cost and COVID-19 positive issues early detection and diagnosis. Regarding this particular disease's novelty, radiological image-based diagnostic methods show some limitations despite their several useful applications in diagnostic centers. This study provides a classification based on various imaging modalities like CT scans and X-ray images. It is shown by some studies that in COVID-19 patients, CT scan had a higher sensitivity. Thus, it has been suggested that a CT scan can assist the COVID-19 diagnosis.

**Keywords** Coronavirus · COVID-19 · CT and X-ray imaging

---

I. Singh (✉)
School of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India
e-mail: Er.inderpal13@gmail.com

P. Chawla (✉)
Department of Computer Science & Engineering, National Institute of Technology, Warangal, Telangana, India
e-mail: priyankachawla.cse@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_30

# 1 Introduction

## *1.1 Background*

A Chinese doctor from Wuhan, China, discovered a coronavirus sickness in December 2019, which quickly spread over the world. The researchers named the virus as CoV because the shape of these viruses was identical to the solar corona when checked under the electron microscopes. In December 2019, it was reported for the first time that the COVID-19 virus had infected a man in China. Since then, the COVID-19 has spread around the globe like wildfire, with a presence in 213 nations [1]. COVID-19 itself is proving to be a complex disease. It is able to emerge in different forms and levels of severity and can be the cause of organ failure and death, ranging from mild to severe, i.e., from light, self-limiting respiratory tract disease to serious progressive pneumonia, breakdown of multi-organs, and death. The rise in this pandemic, the steady increase in the number of confirmed patients, patients who experienced a serious respiratory breakdown and cardiovascular problems are all a big reason to be worried regarding the effects of this virus. Much attention is being given to determine the suitable methods to reach a solution for the problems caused by COVID-19. Even so, the researchers and decision-makers have to face another major problem, i.e., the big data caused by increasing numbers of date during the process of fighting and eliminating this deadly virus [2].

According to a WHO study, the total number of patient afflicted by COVID-19 in the world was 131,837,512 by April 7, 2021, with a death rate of 2,862,664. In light of COVID-19's growing global influence, a variety of initiatives are being made to find remedies to combat its effects. To stop the global pandemic, governments' actions are primarily responsible. For instance, securing a specific area in order to prevent the spread of infection and to ensure that the medical system is capable of dealing with the outburst, delivering the emergency packages to reduce the effects on the economy and people of the nation and adopting adaptable measures in accordance with the COVID-19 situation. On the contrary, research and developments related to COVID-19 are now on priority and have gained specific interest from different stakeholders such as government, industry, and educational institutions. For example, global supply chain, this includes feasibility, consistency, robustness, and strength when considering many supply chains [3]. Apart from this, control and prevention of infection include hand hygiene, keeping a physical distance of at least one meter with others, avoid touching the isolation and informing the recent symptoms to the local medical center. These steps, used combined, are critical for preventing SARS-COV-2 transmission from person to person [4].

The quick breakdown of SARS-CoV-2 has resulted in shortage of clinical devices. Along with the scarcity of masks and ventilators, the testing capacity for the virus has been badly restricted worldwide. Thus, the testing of the suspects and medical personnel was being done on a priority basis. Although to hold the pandemic, thorough testing and diagnosis play an important role. Countries that have been successful in the massive testing of suspects and joint monitoring of virus-affected citizens have succeeded in controlling this SARS-CoV-2 virus. For the majority of the countries, the inadequate capability of testing has led to the demand for alternate techniques to diagnose the COVID-19. Besides, the reliability of the present lab tests. The swab methods as well as the location have a big impact on RTPCR [5]. This global pandemic has put the human population in jeopardy and necessitates immediate action. COVID-19's epidemic has encouraged researchers to provide front-line workers, such as medical personnel, with cutting-edge studies for reduction, diagnosis, and prevention.

The researchers have considered coming up with plans for limiting the emergency, and which can also help in the prevention of such epidemics in future. In addition to health science researchers as well as virology experts, scientists assisted with digital techniques have faced the epidemic with novel techniques. For the battle against COVID-19, two important scientific societies assisted by digital techniques can be recognized. In this regard, the largest digital efforts for the finding of COVID-19 utilizing computed tomography and X-ray pictures have come from the AI community. In addition to these two scientific societies, more attempts are being done to analyze the societal and emotional conduct using social media, collection of scholar articles for the know-how discovery, diagnosis of COVID-19 from collecting the sample of coughing, and automatic contact tracking [6].

### *1.2 Role of AI in Diagnosis of COVID-19*

Machine learning is based on the premise that a system should be able to self-learn and adapt to a variety of real-world occurrences without requiring external programming. Standard programming languages such as Python and R are used in machine learning models for this purpose. The availability of open-source data is the most significant barrier to machine learning application. If all of society's data is available, machine learning technologies can aid in the fight against COVID-19. The machine learning technique is based on the premise of diagnosing COVID-19 with CT scans and X-rays in order to alleviate the load created by a lack of RTPCR test kits. Reports on statistical and epidemiological study of COVID-19 instances can be utilized to determine the relationship between man's movement and viral dissemination. Besides, social media data mining can produce an analysis of sentiment and socio-economic for policymakers in the present epidemic. Hence, to assist the scientific efforts, there is the necessity to collect a new data set related to the man's mobility, epidemiology, behavior as well as radiology.

The realization is important that the pandemic has reached its peak due to which the medical facilities are exhausting. With the increasing number of patients, the emergency services, ICU, have been extended away from their normal facility. In such emergency, there is the need to make quick choices by the medical care providers as well as by the family members of the patient with the little information. The early stages of COVID-19 infection are characterized by little or no symptoms, but then quickly progresses to make the patient exceptionally serious and can lead toward the deadly results due to the failure of multiple organs. The aim is to use this little information and resources to control the sudden degradation and detection of the infection as quickly as possible.

The COVID-19 detection technique based on automated CT scan techniques deals with the training of the learning prototype by the use of existed CT scan data sets that includes the labeled pictures COVID-19-positive and COVID-19-negative cases. Furthermore, samples from both healthy and infected people are required for the diagnosis of COVID-19 from cough in order to understand and distinguish the characteristics of an infected individual from those of a healthy human. As a result, it is critical to provide open-source information and methodologies for the industry.

(a) Researchers worldwide can improve and vary the existed work to restrict the universal pandemic.
(b) Researchers have validated the existed methods for accuracy all over the world before implementing it in real world.
(c) In society-oriented research and development, researchers collaborated to collect data sets and improve the operation of AI/ML algorithms.

### *1.3 COVID-19 Medical Image Processing*

Over few decades, advancement in medical science has greatly altered medical care helping doctors in the identification and treatment of illness more accurately. Doctors like any other human being are vulnerable to errors. A doctor's academic qualifications are determined not just by their intellect level, but also by how they address patients' problems. This mixture appeals to such vast varieties for health outcomes, and in this respect, machine learning is one of the best options for the improvement in the efficiency of a doctor to diagnose and in the treatment of patients. The efficacy of machine learning algorithm relies on different kinds of extracted characteristics and depiction of information. There are two prime concerns faced by machine learning algorithms, i.e., being efficient to scan each data set with high dimensionality and learning of the prototype to identify the suitable task. In predicting the illness precisely, the deep learning method is widely used. In the medical care world, the uses of deep learning methods have proposed new breakthrough. Multiple sources of healthcare information in the actual world comprise MRI, X-ray, positron emission tomography, and CT scan, which has introduced doctors to the massive amount of data. In image processing as well as analysis, the CNN algorithm is preferably used. After the sudden outburst of COVID-19, many studies have been done on processing the information correlated deep learning algorithms, particularly CNN. These studies have mostly focused on using various algorithms and deep learning methods to recognize and discriminate COVID-19.

## 1.4 COVID-19 Radiology Image Analysis Using Deep Learning

Image-based diagnosis approaches have played a critical role in the screening of infected cases during this pandemic. One of the most important radiological procedures for detecting and diagnosing COVID-19 is a positron emission tomography (PET) scan, CXR, or computed tomography scan. Radiological pictures with deep learning are evaluated around two general concepts: "identification" and "diagnosis" for the detection of COVID-19.

### 1.4.1 X-ray Image

For the analysis of COVID-19, the first approach was the X-ray method, and also it has benefits such as cheap as well as has a lower risk from radiation to humans. COVID-19 detection with X-ray is a difficult process. Because the white patches containing water and pus are very long and intricate, the radiologist should pay close attention to detect them.

The medical professionals can access X-ray as a feasible symptomatic procedure for instant and rapid detection of COVID-19 to complement present diagnosis and symptomatic procedures. In the fight against SARS-nCOV-2, a variety of other cutting-edge algorithms are being utilized to improve CXR image results. Because X-rays have a high rate of inaccuracy, CT scans could be utilized to get a more precise diagnosis [7].

### 1.4.2 Computed Tomography (CT) Scan

If comparing a conventional X-ray scan with computed tomography scan, the image acquired by computed tomography can obtain a more accurate and improved image of the chest. But, for patients, CT scan is substantially more expensive as compared to X-ray. During CT scan analysis, numerous slices from every suspected human are provided. To detect COVID-19, physicians and radiologists have to deal with a huge number of CT scan images. The contagion instigates a wide range of computed tomography scan imaging findings, usually ground-glass opacities as well as lung periphery consolidations.

For the analysis of COVID-19, the sensitivity of chest CT has been observed to be considerably higher as well as it can happen earlier a positive viral lab test. As a result, in pandemic areas where the basic medical system is under strain, hospitals with large admissions use computed tomography for fast assessment of patients with possible COVID-19 sickness.

Chest CT is important in estimating COVID-19 patients with significant and complicated respiratory symptoms. On the bases of these computed tomography scan images, it can be possibly assesses whether the lungs are affected severely, and

**Fig. 1** Day 5 after COVID-19 symptom, day 15 after COVID-19 symptom, and day 20 after COVID-19 symptom

how the disease of a person is progressing which will be useful in making healthcare choices. Figure 1 indicated few samples of CT scan images:

Although, the two words, i.e., diagnosis and detection, are being used indifferently in the world of healthcare image analysis. In healthcare textbooks, finding is believed to be an introduction to diagnosis. Likewise, these two words vary medically in the context of COVID-19 where detection is defined as the COVID-19 infected cases over non-COVID-19. This implies that for the non-COVID-19 patients, there is no available data about the disorders; this population can have a distinct variety of bacterial pneumonia, viral, or any other group of coronavirus disorder but no COVID-19. The term diagnosis is defined to separate COVID-19 from other contagious lung disorders, for instance, different kinds of lungs diseases. The term diagnosis is significant for the groups where the other disorders are predetermined, and COVID-19 can be differentiated certainly from different kinds of pneumonia or other coronaviruses. Most of the researchers and doctors have relied on simple radiographic images or X-rays for COVID-19 diagnosis. Still, these images lacked the required resolution to diagnose the COVID-19 accurately and have many drawbacks.

This is why artificial intelligence researchers used deep learning as an efficient tool to assist medical specialists to enhance the diagnostic of COVID-19. Based on the nature of DL in extracting the characteristics of the images, the said technology is efficient in the detection of COVID-19 patient as well as in the extraction of infected lung tissues; a lot of studies have used different types of DL methods for the analysis of these images. Although there is advancement in the detection and diagnosis of COVID-19 using deep learning algorithms, the major disadvantage of this method is the shortage of required devices or tools for the diagnosis of COVID-19 throughout the healthcare and diagnostic centers. Moreover, several COVID-19 patients need many CT scan images of their chest. Radiation exposure to the patients can cause severe issues during computed tomography scan.

### *1.5 Research Aim*

The loss of life was the most devastating aspect of the epidemic, [6, 8] but the implications in the aftermath were just as bad: survivors' psychological well-being and the socio-economic fallout were both extremely unpleasant. Early detection and isolation of patients are the most effective strategy to stop the spread of COVID-19. The project's major goal is to detect the COVID-19 virus early on in order to decrease human-to-human transmission, enhance testing rates, and deliver accurate results. A number of methods are available that can detect COVID-19 at early stages and prevent the loss of life. Out of all the technique available, the main focus of this paper is on DL because of its self-learning ability in which CT scan is used to detect the COVID-19 in a fast and effective way. A literature survey is conducted in which a number of techniques proposed by various researchers are reviewed and analyzed. These techniques had some drawbacks such as poor image quality, high error rate, time consuming, etc., that need to be addressed. The motivation behind this study is to reduce or eliminate the above-mentioned problems of traditional techniques. By eliminating these problems, the accuracy and efficiency can be achieved in early detection of COVID-19, which ultimately will increase the testing speed and can reduce the transmission of COVID-19.

## 2 Literature Survey

In this section, related work and research gaps have been discussed. Table 1 shows the related work in the field.

### *2.1 Research Gaps*

1. Images obtained from X-rays machines, CT scans, etc., mostly get affected due to machine noise, which gives blurry vision and loss of information in the image. This could lead the degradation of models and application those are utilizing the feature for any specific task. Very few researches have been done to eliminate these noises from images.
2. Feature extractions are key step for any classification or prediction model in any research domain. Efficient features help in understanding the patterns and improve the detection rate. In COVID-19 detection system especially where medical images are used to detect the disease, feature extraction models are playing important role for early detection. Current feature extraction models are mostly in transformation-based. Further work can be done in improving these techniques by collaborating the features of multiple models.

**Table 1** Brief study of various current COVID-19 detection systems

| Author’s name | Publishing year | Description |
|---|---|---|
| Ciotti et al. [9] | 2020 | Suggested a novel 3D CNN to concentrate on in-fiction areas in the lungs while making decisions regarding COVID-19 diagnosis |
| Shuja et al. [1] | 2020 | Suggested new method COVID-19 lung infection segmentation based on DL to detect automatically the factoring in chest CT images obtained |
| Arellano and Ramos [2] | 2020 | Developed a model that was based on weakly supervised DL method in order to identify and categorize COVID-19 from CT pictures |
| Ozsahin et al. [3] | 2020 | Used the HMM as well as the Viterbi algorithm and recommended an efficient unsupervised technique for identifying and localizing the pleural line in LUS images |
| Singh et al. [4] | 2020 | They suggested a reliable method to detect COVID-19 pneumonia through digital CXR images by applying pre-trained deep learning techniques, thus optimizing detection rate |
| Imran et al. [5] | 2020 | When one class unexpectedly outweighs the other, an activation function based on GEV distribution outperforms the traditional sigmoid activation function |
| Elsayed Abd Elaziz et al. [6] | 2020 | Proposed a unique joint learning technique to identify COVID-19 in which heterogeneous data sets along with distribution discrepancy for learning purpose were used |
| Wang et al. [8] | 2020 | Suggested a chest radiograph classification (DL-CRC) method based on effective DL technique to precisely make different corona cases from other normal and abnormal cases |
| Li et al. [7] | 2020 | The authors demonstrated that a set of DL-based algorithms might be trained on a broad multinational cohort of patients to optimize parietal lung parenchyma together with COVID-19 pneumonia categorization |
| Hussain et al. [10] | 2020 | Tackled the cough issue in the respiratory system by comparing alterations of path morphological that were caused by the infections of COVID-19 while comparing other respiratory infections |

(continued)

**Table 1** (continued)

| Author's name | Publishing year | Description |
|---|---|---|
| Mishra et al. [11] | 2020 | This paper proposed a new technique based on machine learning to identify images of chest X-rays in two subgroups: non-COVID-19 people and COVID-19 patients. Fractional multichannel exponent is a new fractional component. Characteristics from chest X-ray pictures were extracted using moments |
| Patel et al. [12] | 2020 | To fight against COVID-19, the major focus was on regions that utilized cloud computing oriented with artificial intelligence to overcome challenges that were faced by various researchers and practitioners while dealing with big data of COVID-19 |
| Dash et al. [13] | 2020 | The authors of this paper utilized an AI-based algorithm to combine clinical symptoms with chest computed tomography findings, infection history, and laboratory testing. 46.3% of 905 patients were COVID-19 tested positive |

3. With the use of classifiers such as GAN, LTMS, and ELM, the accuracy of detecting COVID-19 at early stages is almost 80%. Therefore, using or developing an advanced classifier in future, techniques are recommended for achieving high accuracy.
4. The multimode techniques for extracting features from images at various levels yield high accuracy in detecting COVID-19. However, not much work has been done in this field.
5. Very few studies are available on classifying different stages of COVID-19; most of the works are only on detection either the suspect is patient or normal. If stages-oriented models can be designed, the medical treatment or related measures at those stages can save the patient for getting severe disease.
6. It has been seen that recently, researchers have focused on the deep learning, as these systems were able to detect COVID-19 automatically in fast and effective way. Therefore, analyzing current DL techniques and improving their performance may help in detecting COVID-19 at early stages that can save human life.

## 3 Proposed Work

The main contribution of the proposed research will be focused on providing a classifier and detection model for detecting COVID-19 in early stages, so that the necessary diagnostic measure can be taken, and life of the patient can be saved. To achieve this, proposed model would focus on three different phases of detection approach. The very first phase will be working on proposing a pre-processing model for medical images.

As using various machines, it captures the medical images; therefore, these images get affected by noise or some light related issues. Therefore, the proposed model will be capable to handle these issues and refine the image quality before passing it to leaning network.

The next phase of the research will be focused on providing a segmentation approach that can easily extract the virus-affected region from whole image. Finally, the proposed study will provide a deep learning algorithm's architecture that is expected to be improved with respect to currently available systems.

The suggested method's layers and feature learning strategy will be designed in such a way that the COVID-19 can be detected quickly and accurately.

Other main contribution of the proposed study will be focused on classifying the COVID-19 at different stages; most of the current systems are only detecting the positive or negative class of the input. But in proposed model, various stages of the COVID-19 will also be taken into consideration so that the patient can be detected in early stages, and their life can be saved. The tentative design of the model is as shown in Fig. 2: The suggested COVID-19 detection system's tentative architecture diagram is shown in Fig. 2.

1. In the first stage, the images are gathered from the data sets that are available online on few government official sites or research Websites (Example, Kaggle, UCI machine learning, etc.)
2. These images are then analyzed and processed in the next stage in order to remove the unwanted noises present in the image, which produces blurred vision. The contrast of the images is can also be enhanced by using the filtering techniques to increase image quality.

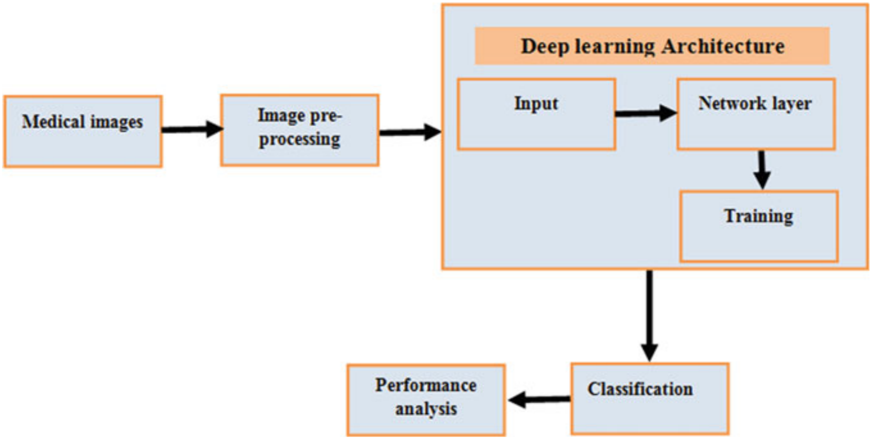


**Fig. 2** Block diagram of COVID-9 detection in DL system

3. The images are then passed to the proposed deep learning architecture in which the enhanced images act as an input for the system. These images are then processed at different layers of the network includes feature extraction, weights generation (in such a way that maximum information can be obtained from the processed image), training of network, etc.
4. The model trains itself with the information obtained from the images. Finally, classification is done in which the proposed system will detect whether the person is COVID positive or not.
5. On the basis of this detection, the performance of the model will be analyzed to validate its efficiency in determining COVID-19 at early stages.

## 4 Conclusion

Since the pandemic, COVID-19 has been widely spread throughout the world, resulting in an exponential increase in the number of cases. To avoid the healthcare system becoming overburdened, a method that aids in disease diagnosis in individuals must be developed, and it must be both quick and inexpensive. For radiologists, employing these imaging modalities approaches in COVID-19 diagnostic decision-making is a key tool for reducing human error as well as assisting them in making choices at disease's peak and in critical conditions.

## References

1. Shuja J, Alanazi E, Alasmary W, Alashaikh A (2020) COVID-19 open source data sets: a comprehensive survey. Appl Intell 1–30
2. Arellano MC, Ramos OE (2020) Deep learning model to identify COVID-19 cases from chest radiographs. In: 2020 IEEE XXVII international conference on electronics, electrical engineering and computing (INTERCON), Lima, Peru, pp 1–4
3. Ozsahin I, Sekeroglu B, Musa M, Mustapha M, Ozsahin UD (2020) Review on Diagnosis of COVID-19 from chest CT images using artificial intelligence. Comput Math Methods Med 1–10
4. Singh D, Kumar V, Kaur M, Kumari R (2022) Early diagnosis of COVID-19 patients using deep learning-based deep forest model. J Exp Theor Artif Intell 1–11
5. Imran A, Posokhova I, Qureshi HN, Masood U, Riaz MS, Ali K, John CN, Hussain MDI, Nabeel M (2020) AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. Inf Med Unlocked 20:100378. ISSN: 2352-9148
6. Elsayed Abd Elaziz M, Hosny K, Salah A, Darwish M, Lu S, Talaat A (2020) New machine learning method for image-based diagnosis of COVID-19
7. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J (2020) Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology 296
8. Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, Wang M, Qiu X, Li H, Yu H, Gong W, Bai Y, Li L, Yongbei Z, Wang L, Tian J (2020) A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. Eur Respir J

9. Ciotti M, Ciccozzi M, Terrinoni A, Jiang WC, Wang CB, Bernardini S (2020) The COVID-19 pandemic. Crit Rev Clin Lab Sci 57(6): 365–388
10. Hussain A, Bouachir O, Al-Turjman F, Aloqaily M (2020) AI techniques for COVID-19. IEEE Access 8:128776–128795
11. Mishra AK, Patel A, Jain S (2021) Impact of covid-19 outbreak on performance of indian banking sector. In: CEUR workshop proceedings, vol 2786, pp 368–375
12. Patel A, Debnath NC, Mishra AK, Jain S (2021) Covid19-IBO: a Covid-19 impact on Indian banking ontology along with an efficient schema matching approach. N Gener Comput 39(3):647–676
13. Dash S, Chakravarty S, Mohanty SN, Pattanaik CR, Jain S (2021) A deep learning method to forecast COVID-19 outbreak. N Gener Comput 39(3):515–539

# Optimal Convolutional Neural Network Model for Early Detection of Lung Cancer on CT Images

**Dugyala Vignesh kumar, Guttula Nookaraju, T. Sreeja Reddy, B. S. Panigrahi, and S. N. Mohanty**

**Abstract** Lung cancer diagnosis and prediction in the earliest stage of the disease can be extremely helpful to improve the survival rate of patients, but diagnosing cancer is one of the most challenging tasks for a radiologist. In this paper, we propose that convolutional neural networks (CNNs) can effectively detect and predict lung cancer. The design of model has 8 layers in which one is input layer and two conv layers and a ReLU, max-pooling to extract features and a fully connected layer which has sigmoid activation function which in terms connected to 1 neuron which is an output layer. This ouput layer is responsible for classifying the image in to a benign or malignant nodules. We have acquired the dataset from Kaggle having 1097 cases which can be passed as input data for input layer. A 2D set of resampled images with 64 * 64 pixels of size which are rotated, scaled, and randomly translated images is generated as input samples.

**Keywords** Convolutional neural networks (CNNs) · Max-pooling · Rectified linear unit (ReLU)

## 1 Introduction and Background

Cancer is a condition when cells grow out of control. Cells in any part of the body are prone to cancer and have the tendency to easily spread to surrounding tissues and organs. Lung cancer is generally recognized with name carcinoma. It is defined by a peculiar developement in cells. Carcinoma is the cause of high cancer-related death rate. This process can be unqualified, but eventually spontaneous to deliberate.

---

D. Vignesh kumar · G. Nookaraju (✉) · T. Sreeja Reddy · B. S. Panigrahi · S. N. Mohanty
Vardhaman College of Engineering, Hyderabad, Telangana, India
e-mail: rajuguttula1798@gmail.com

D. Vignesh kumar
e-mail: vigneshdugyala980@gmail.com

B. S. Panigrahi
e-mail: bspanigrahi@vardhaman.org

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*, Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_31

When left untreated or undiagnosed, cancer can become deadly by time and might be fatal to the person.A carcinoma is a type of cancer that begins in the cells that line the skin or an organ such as the lungs or kidneys. Cancers of the lungs are classified as small-cell lung carcinoma or non small-cell lung carcinoma.

Cancer is generally a genetic disease, which does not exclude it from being caused by external affiliations. It is more in relation to the changes that are associated to the gene of the body that might even be caused due to certain stimulations the body responds to, which might alter the cell behavior, especially how they divide. However, most lung cancers are not genetically inherited diseases, although there is a genetic component determining this pathological tendency. The primary cause of lung cancer is smoking. The occurrence of cancer in a few cases was found in people with no smoking history but with symptoms of exposure to air pollution, secondary smoking, and some toxic gases. Before, the twelfth-century occurrence of lung cancer was very rare. But comparing the stats, it is usually caused due to exposure to polutant chemical air, in many forms. Like any other cancer, early stage diagnosis is very important to treat lung cancer. The X-ray image of lungs can show the abnormal growth of cells in lungs due to cancer. Cancer can also be detected by sputum cytology, a method of analyzing the subject's sputum to identify the affected cells. Biopsy is another method of collecting a tissue sample from the suspected area and running tests to determine the possibility. Each of these methods have different accuracy rates, based on the stage of cancer.

In conventional method of detection of lung cancer, we are using radio graph techniques and computed tomography (CT). We can achieve better results using CTscan compared to a normal X-ray because in CTscan, we use a computer and an X-ray to visualize the lungs. Using CTscan, doctors can visualize the lungs in different angles since it gives the clear image compared to a normal one. In this study, 20 lungs images samples have been taken into consideration. Thus, the study could help in the early detection of lung cancer with greater accuracy.

## *1.1 Literature Survey*

The project proposed here uses an ML model with convolutional neural networks that emphasize the method of utilizing filtering features to initialize a cognitive methodology for faster decision-making and accuracy. Several other methods involving but not limiting to identification of the case follow similar approach with a little or no directive process. Predicting optimal subset of genes is a systematic procedure of analyzing gene microarray of expression data to predict the most probable cancer causing agent. While several classification techniques like multi-layer perceptron and random subspace are compatible with various subset of attributes, the precision recall value of sequential minimal optimization (SMO) is much better. Also, support vector machine (SVM) classifier is used for multi-stage classification for the detection. However, other methods like electrical impedence tomography proved to be superficially better at understanding the cancer development by visualizing the respiratory system. Around the exact idea of promoting the salient feature extraction and classification, though biomedical methods have near field monitoring and direct analysis, using CNN and machine learning to analyze the possibility of cancer development can be considered as an early stage of developing a sustained diagnosis. Other existing procedures are time consuming, and relying solely on them is waling on a thin line especially where time is a deciding factor. CNN algorithm works on various layers of classification which involves various criteria like max-pooling and ReLU over various stages. While biopsy methods are still proven to be most used, deep learning models are have high success rate since they work in a hierarchy. Descriptors like GLCM, MPLNN, and DT help in propagating the further ML structure to maximum accuracy possible after filtering. Cancer prediction using CNN follows image processing with segmentation and multiple channeling where the data is used to train the model, and then, the features are extracted using CNN. The extracted data is compared and validated after classification, thus predicting the cancer.

The factor of achieving the required verification through deep prognosis is different in various methods of justification and imaging techniques. Although certain demographics suggest the importance of the area of analysis using MRI or CT, the concrete results are always varied based on the topology. Studies show that sensitivity of positive prediction through CT and MRI is exclusively higher than biopsies, thus identifying the structural value with image sourcing through CT helps the classifier machine acquire high initial predictive discrimination followed up by constructive feature extraction for accurate classification. The different image processing techniques employed through GLCM after de-noising takes minimum time to produce the result (Figs. 1 and 2).

## 2 Proposed Model and Method

### *2.1 Model Development Using CNN*

In our study, we are using convolutional neural network algorithm which applies gradient descent for training the model [1, 2]. CNNs are widely used for classifying two-dimensional structure like images and that is why CNNs are treated as most powerful algorithms. Mostly, CNN algorithms are called as end-to-end algorithms which contains various sub-sampling, pooling and convolutional layers which are all connected to a final layer which is fully connected [3]. Firstly images of size of $64 \times 64$ pixels are given as input for the input layer. We can see the architecture of CNN and training process of the algorithm in Fig. 3. The CNN architecture has 6 layers in total: The primary layer is the input layer, the last layer is called as output layer, layer two and four are called as convolutional layers, and layer three and layer five are sub-sampling layers made of max-pooling layers and ReLU activation [4]. We use batch normalization for feature extraction, and the output layer is made up of sigmoid activation which is connected to one neuron. This output layer classifies

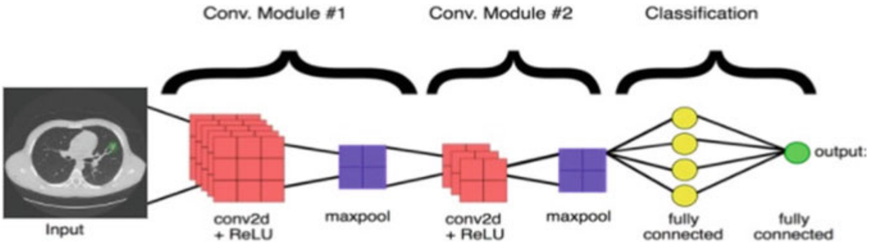


**Fig. 1** Architecture of CNN algorithm

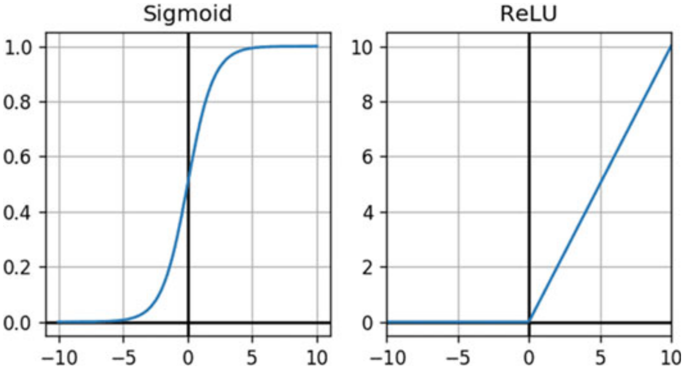


**Fig. 2** Activation diagram

the image in to benign tumor or malignant tumor which are the categories of classes available to classify. In convolution layers, we extract the features by preserving spacial relationships of the pixels of the input images which in turn reduces the size of the images and at the same time preserving the features of the images which helps in reducing the computation cost. To be most specific on the design of the CNN, the 2nd layer has 32 feature maps with 3 × 3 dimensions, which is followed by max-pooling layer of 2 × 2 pixels [5, 6]. Down-sampling for the convolved image is performed by the pooling layer. With the help of max-pooling layer, we can reduce the dimensions of the image by preserving the features which in turn reduces the computational cost. And at last, the output layer has 1 × 1 matrices with sigmoid function having which has one neuron and that is responsible for matching the output to the available categories (benign, malignant). In turn, each and every neuron receives linear combinations of input from its corresponding neurons, with weights and biases based on the prior layer's input. Finally, the output layer computes the probability of strength based on the data from the prior layer.

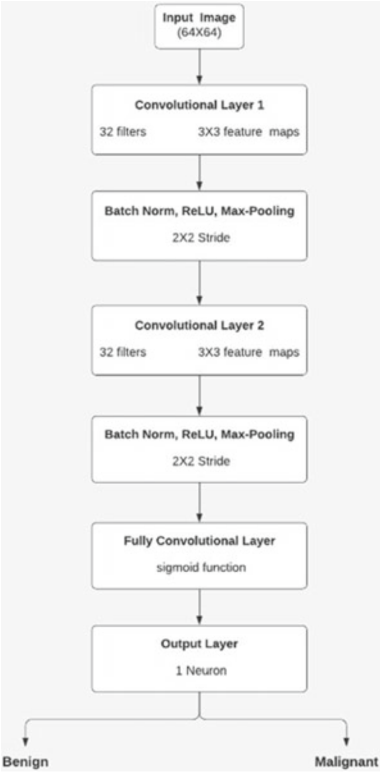


**Fig. 3** Training process of CNN

## 2.2 *Methodology*

**Dataset Collection**

We have collected the dataset from Kaagle. The data in the dataset is gathered from a hospital with in a span of 3 months. This dataset contains the lung CT scan images of different patients which were diagnoised from lung cancer in benign or malignant stages. This dataset has a total of 1100 CT scan images which are from slices of 105 cases. The cases are divided into two categories: benign and malignant. Thirty five cases are malignant, and seventy are benign. We de-identified all images before analyzing them. A scan consists of several slices. There is a range of 70–110 slices, with each slice representing a different side or angle of a human chest.

**Preprocessing of Image**

In the process of training a model, the first step we do is preprocessing. It is like cleaning the data such that there will not be any corner cases which breaks our model. Preprocessing makes our model robust by making it to accept the data of any dimensionality. Few preprocessing techniques we have implemented are checking for the false data and remove it, resizing and re-scaling the images to the desired dimensions, and data augmentation for making the images of different views.

**Image Resizing**
Every image which is given as an input image to the convolutional neural network must be resized to a specific dimensionality which can be done using image resizing. Image resizing is one of the crucial preprocessing techniques. We can resize the images in two different approaches by down-scaling the images and by cropping the borders. In the second approach, we may end up losing the border data which can become the crucial features of the image. In the first approach also, we may end up getting the deformed images. But deformed images are mostly a reasonable choice than cropped images. So, the first approach is more feasible for most of the applications. Here, we get the deformed images, and we are not at the risk of loosing the patterns or border features. Here, we are performing re-scaling and resizing to make input image match with the desired dimensions of model.

**Data Augmentation**
To avoid overfitting, CNN requires large data. Overfitting happens when a model gives better accuracy on training data but gives least accuracy on test data. This may happen due to less quantity of training data. Data augmentation is a technique which prevents overfitting, and it helps in generating new lung CT images from the data what we have by applying certain mathematical functions. In this model, every image is transformed using zooming, rotation, and flipping. Horizontal axis flipping is mostly used than vertical axis flipping. Usually, augmentation through rotation is implemented by making the image rotate either left or right. We have implemented the augmentation process with the help of the image generator of keras library.

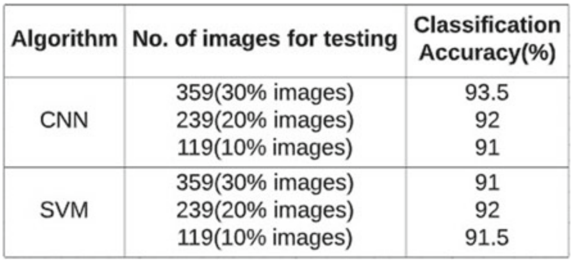
## 3 Results

We made comparison between CNN and SVM algorithms. And we got the accuracy difference of 2.5, CNN being the better one among two. We used 30, 20, and 10% of images from the proposed dataset of 1100 images for testing purpose.

## 4 Scope

We made some observation to support future work

1. Though the end results, we acquired on the study were good, as we focused on limited data with restricted count of algorithms. Hence, we can increase the scope with a large volume of data could help in elevating the performance.



| Algorithm | No. of images for testing | Classification Accuracy(%) |
|---|---|---|
| CNN | 359(30% images) | 93.5 |
| | 239(20% images) | 92 |
| | 119(10% images) | 91 |
| SVM | 359(30% images) | 91 |
| | 239(20% images) | 92 |
| | 119(10% images) | 91.5 |

**Fig. 4** Accuracy comparision between different model based on the number testing images

2. A rather radical investigation on optimal input dimensions in deep learning models can yield better accuracy.
3. Broadening the input dimension to 3D data feature and building the relevant CNN model (Fig. 4).

## References

1. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. https://doi.org/10.1109/5.726791
2. Tekade R, Rajeswari K (2018) Lung cancer detection and classification using deep learning (ICCUBEA). https://doi.org/10.1109/ICCUBEA.2018.8697352
3. Tafti AP, Bashiri FS, LaRose E, Peissig P (2018) Diagnostic classification of lung CT images using deep 3D multi-scale convolutional neural network (ICHI). https://doi.org/10.1109/ICHI.2018.00078
4. Neal Joshua ES, Bhattacharyya D, Chakkravarthy M, Byun YC (2021) 3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation. https://doi.org/10.1155/2021/6695518
5. Yu J, Yang B, Wang J, Leader J, Wilson D, Pu J (2020) 2D CNN versus 3D CNN for false-positive reduction in lung cancer screening. https://doi.org/10.1117/1.JMI.7.5.051202
6. Chon A, Balachandar N (2017) Deep convolutional neural networks for lung cancer detection

# Lung Cancer Diagnosis Using Deep Convolutional Neural Network

**Jasdeep Kaur Basra, Eshaan Gupta, Munish Saini, and Amit Chhabra**

**Abstract** Lung cancer is a malignant tumor of the lungs that causes uncontrolled cell development in the tissues, and it is the world's most lethal cancer. The main aim of our study is to predict lung cancer type and its severity using a two-dimensional convolutional neural network (CNN). For achieving the objective, computed tomography scans of different types of lung cancer (adenocarcinoma, large cell carcinoma, squamous cell carcinoma) and healthy lungs were collected. Further, the improvisation of the CT scans was done using the Weiner filter. The improved images were fed into the CNN model. Flask app is used to anticipate the type and the severity of lung cancer for the validated images. The accuracy achieved by the CNN model is above 90%. The accuracy achieved for different classes such as adenocarcinoma, large cell carcinoma, squamous cell carcinoma, and healthy lungs is 90%, 92%, 91%, and 92% respectively. Moreover, other metrics are included to evaluate our model which involves precision, recall-measure scores, and the results are calculated as 0.99, 0.90, and 0.94 respectively. The usage of the flask app in prediction makes our study real-time monitoring. Moreover, this study will help doctors in the diagnosis of lung cancer in a fraction of seconds due to the usage of the flask app.

**Keywords** Lung cancer · Convolutional neural network (CNN) · Prediction · Image improvisation

## 1 Introduction

A neoplasm or cancer is an irregular tissue mass that exceeds the growth of normal tissue and is not synchronized with that of normal tissue and continues in the same unrestricted manner after the factors that caused the change have ceased.[1] As reported by the WHO, the most common cause of mortality is cancer in which liver, lung,

---

[1] https://www.news-medical.net/health/Neoplasm-Definition.aspx.

J. K. Basra (✉) · E. Gupta · M. Saini · A. Chhabra
Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, Punjab, India
e-mail: malhip78@gmail.com

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_32

stomach, colon, and breast cancer are some of the most common among the organisms. It is mainly caused due to excessive cigarette smoking.[2] One of the leading causes of death is lung cancer among males and females in the USA.[3] Lung cancer is mainly of two parts which consist of non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is sub-categorized into three main parts such as adenocarcinoma, squamous cell carcinoma, and large cell carcinoma[2]. Lung cancer can be contained if the proper diagnosis is done at the early stages, and the type of cancer is known beforehand. The present method of detection includes computed tomography (CT) scanning of the chest which is the foundation of lung cancer imaging, based on which further management is determined.[4] CT scan uses X-rays making a detailed cross-sectional image of the body. It shows the type, size of any type of lung cancer. This test also reveals the areas such as the adrenal gland, liver, and brain which are affected due to the lung cancer spread.[5]

Traditionally, algorithms for image processing have been used to extract features from images to distinguish between malignant and benign tumors [1]. For distinction between malignant and benign tumors, a development of hand-designed features is required from the images. In recent times, a deep convolutional network is used for the analysis and classification of CT images which gives really promising results in predicting lung cancer as it extracts features from the images automatically.

In this work, we have presented a two-dimensional CNN model to predict the type and severity of lung cancer. To achieve our objective, 1200 CT scan images of non-identical types of lung cancer were collected from the Kaggle and Google images (freely available with common creative licensing).[6] Further, the prediction is made using the flask app that is a Web application framework written in Python in which the validated images are applied, and the prediction is made regarding the type and the severity of the lung cancer. The usage of the flask app makes our study real-time monitoring. Further, image improvisation is also dealt with in this work which helps in achieving better results.

The main objectives of our work are summarized below:

1. To refine the quality of CT scans images depicting different types of lung cancer.
2. To classify lung cancer and its severity using a two-dimensional CNN.

The remaining paper is structured into various sections as follows. Section 2 presents the literature review done in the context of this domain. Section 3 describes the detailed methodology opted to predict lung cancer. Sections 4 and 5 provide the results and discussions, respectively, on the findings of the conducted study. We conclude the paper with some ideas for future work in Sect. 6.

---

[2] https://www.medicinenet.com/lung_cancer/article.html.

[3] https://www.cancer.org/cancer/lung-cancer/about/what-is.html.

[4] https://www.mayoclinic.org/diseases-conditions/lung-cancer/diagnosis-treatment/drc-20374627.

[5] https://www.cancer.org/cancer/lung-cancer/detection-diagnosis-staging/how-diagnosed.html.

[6] https://www.kaggle.com/mohamedhanyyy/chest-ctscan-images.

## 2 Literature Review

Various systems have been implemented to detect various kinds of cancer, but still, many of them have loopholes. One of the researchers has created an automated brain tumor classification approach using full-slide digital pathology images [2]. Moreover, a study investigated automatic classification from histopathological images of breast cancer [3].

Furthermore, there exists a plethora of work being done to classify lung cancer using machine learning algorithms. One of the studies has identified features from chest X-ray images and used the notion of the neural network process of backpropagation to enhance the accuracy [4].

One of the research implemented several machine learning methods on various cancer-related data and deduced that the implementation of feature selection and classification integration would deliver a promising result in cancer data analysis [5]. Furthermore, the support vector machine (SVM) classification method is used to detect nodules in lungs at an early stage [6]. Another study combined non-identical computer-aided methods to identify the nodules in lungs using CT scans [7]. A researcher has incorporated supervised learning such as a linear regression model, vector machine support, and ANN and deduced that compared to other approaches, SVM results in better accuracy of 96% [8]. The main drawback of these techniques is that they need manual extraction of the feature, whereas the usage of CNN can lead to automatic feature extraction.

Moreover, another study conducted a survey on the significance of the convolutional neural network for lung module prediction with almost 90% accuracy [9]. A computer-aided diagnostic procedure has been used in the study in which the process of segmentation, identification, and staging for the categorization of CT lung cancer images with greater precision is followed in three steps [10]. Moreover, research indicated that C4.5 performs better with greater accuracy on the North Central Cancer Treatment Group (NCCTG) lung cancer data and also predicted that C4.5 is a better classifier with increased data on lung cancer training data [11].

In one of the studies, the researcher developed an algorithm based on CNN and used staging fluorodeoxyglucose positron emission tomography (FDG-PET)/CT pictures for differentiating lung cancer lesions as T1-T2 or T3-T4 [12]. The accuracy of thee model resulted as 87%. Another study has used a 3D convolutional neural network to automatically differentiate lung cancer, and the accuracy achieved in the model was 69.7% [13].

A 3D CNN model has been used on data carrying details about the patients suffering from lung cancer to identify nodules and also detects interest points using U-Net architecture. The accuracy achieved by their model was 86.6% [14].

The studies mentioned above are using various machine learning algorithms which include Naïve Bayes, SVM, linear regression to detect lung cancer, but the main drawback of these methods it performs feature extraction manually which is overcome by our study by using a convolutional neural network that performs feature extraction from images automatically. Further, other studies have also used 3D CNN [14] for prediction of the lung cancer, but the main loophole of these studies is that it only detects the nodules, whereas our research also specifies the lung cancer type and its severity. Furthermore, our study is real-time monitored due to the usage of the flask app as gives users the facility to upload the lung cancer CT scans and predict the result.

Moreover, one of the studies [15] histology was predicted in early stage NSCLC patients using CNN. The main concern of this study was to construct strong deep learning-based radiomics models to help identify clinically significant histologic subtypes in NSCLC without using invasive methods. The CNN was able to identify tumor histology with AUC up to 0.71.

# 3 Methods

This section describes the data analysis and methodology (Fig. 1) opted to predict lung cancer types and their severity.

## 3.1 Data Collection

For carrying out study, 1200 CT scans of lung cancer are collected from Kaggle. Moreover, the retrieved dataset is open to access and can be downloaded freely. Specifically, the dataset consisting of images targeting three types of chest cancers: (a) adenocarcinoma, (b) large cell carcinoma, (c) squamous cell carcinoma, and 1

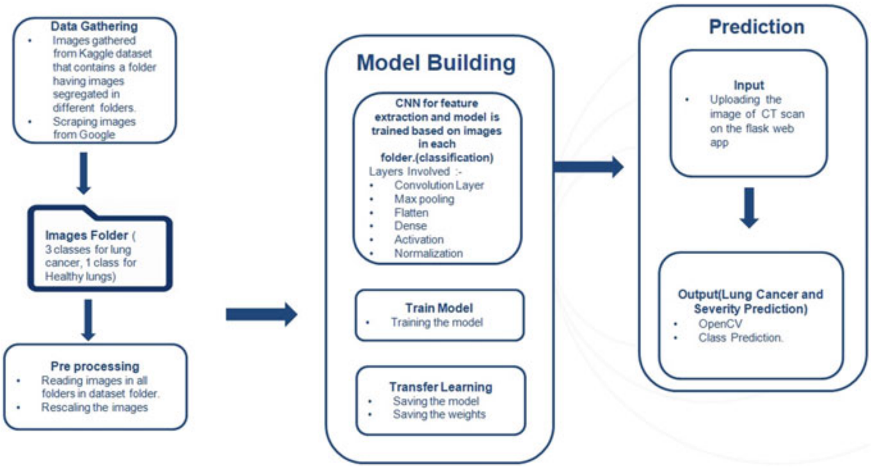


**Fig. 1** Data analysis and methodology

folder for the normal cell. Further, dataset is divided into 70% training set, 20% testing set, and 10% validation set.

### 3.2 Pre-processing

The pre-processing of the images is performed using the image restoration technique. Image restoration removes noise and distorted images degraded by various atmospheric defects. The technique opted in our study to cancel the disruption from the images is the Weiner filter. This filter was presented by Norbert Wiener in 1940 and published in 1949. The main advantage of the Weiner filter is that it removes additive noise and inverts blurriness simultaneously. The Wiener filter reduces the mean square error between the anticipated process and the designated process. In the inverse filtering and noise smoothing method, it minimizes the total mean square error. Wiener filtering approximates the original image linearly. [16]. It is dependent on statistical analysis.

The Wiener filters are distinguished by the following factors as discussed below:

1. Assumption: Stationary linear stochastic imaging and noise processes with known spectral properties, as well as known autocorrelation and cross-correlation
2. Requirement: Physically, the filter must be feasible.
3. Performance criterion minimum mean square error (MMSE): It is commonly used in the deconvolution method [17]. It is also common in use due to its speed and simplicity. It is simple due to the usage of the linear equation in order to determine the collection of optimum filter weights. Furthermore, it predicts cross-correlation and covariance matrixes of noisy signals to quantify these weights and produce a reasonable estimate of the unfiltered deterministic signal in the presence of Gaussian noise. Afterward, the noise statistics are calculated which are further used to calculate the optimum filter weights. This approach is efficient when the distribution of noise is Gaussian. Moreover, its implementation requires only a few computational steps that are very easy to process [18].

Weiner filter in the Fourier domain can be expressed as follows [19]:

$$G(u, v) = \frac{H * (x, y) P_s(x, y)}{|H(x, y)|^2 P_s(x, y) + P_n(x, y)}$$

Dividing the equation by $P_s$ makes the

$$G(u, v) = \frac{H * (x, y)}{|H(x, y)|^2 + \frac{P_n(x,y)}{P_s(x,y)}}$$

where

$H(x, y)$ = Degradation function.

$H^*(x, y)$ = Complex of degradation function.
$P_n(x, y)$ = Power spectral density of noise.
$P_s(x, y)$ = Power spectral density of ungraded image.

The Wiener filter is used to eliminate the noise from the corrupted image found based on statistical inference made from the local region of every single pixel. The noise power affects this filter. When the variance is high, the filter does less smoothing, and when the variance is small, the filter does more smoothing.

## *3.3 Image Data Augmentation*

It is a technique for creating new training data from current training data.[7] It is executed using domain-specific algorithms that produce fresh and diverse training examples from existing training data. Deep learning algorithms nowadays can learn properties that are independent of their location in the image, such as the convolutional neural network, or CNN [20]. However, in this transform invariant learning strategy, augmentation will aid the model in learning features that are likewise invariant to transformations such as left-to-right to top-to-bottom ordering, light levels in images, and more. Image data enhancement is usually limited to the training dataset, not the validation or test dataset. This is different from data preparation like picture resizing and pixel scaling; all datasets that interface with the model must be done uniformly. Keras library is used to perform Image augmentation under which ImageDataGenerator class is implemented.

Data augmentation is implemented using a geometric transformation technique under which the following considerations are to be implemented:

- Flipping: It is a rotating of the image horizontally or vertically. Vertical axis flipping is far less common than horizontal axis flipping. In datasets such as ImageNet, this enhancement is one of the easiest to apply and has shown to be beneficial.
- Color space: A dimension tensor is used to represent digital picture data (channels of height * width * color). Another technique that is very realistic to execute is to perform augmentations in the color channel space. Quite basic color increases involve single-channel isolation of colors such as $R$, $G$, or $B$. By isolating the matrix and adding 2 zero matrices, from different color channels, a picture may be simply turned into its representation in one color channel. Also, manipulation of the RGB values can be easily done to boost or decrease the brightness of the picture with simple matrix operations.
- Cropping: When working with image data that has a combination of height and width measurements, cropping images can be used as a realistic processing phase by cutting each image's center portion [21]. Also, to have a somewhat similar effect to translations, random cropping can be used. The only difference between

---

[7] https://en.wikipedia.org/wiki/Data_augmentation.

random cropping and translating is the input size, such as (256, 256) → (224, 224), would be decreased by cropping, while translations retain the image's spatial dimensions.

- Rotation: By rotating the picture right or left on an axis between 1 and 359°, rotation augmentations are performed. The protection of rotation augmentations is strongly determined by the parameter of the degree of rotation. For digital recognition tasks, minor rotations such as between 1 and 20 or − 1 and − 20 may be useful, but as the degree of rotation increases, the label of the data is no longer retained after transformation.
- Translation: Moving images left, right, up, or down is a highly important transformation for avoiding positional bias in data. For instance, if all of the images in a dataset are perfectly centered, as is frequent in face recognition datasets, the model must also be evaluated on perfectly centered images. Because the original picture is only read in one direction, the remaining space can be filled with either a fixed value, such as 0 s or 255 s, or random or Gaussian noise.
- Noise injection: It comprises injecting a random value matrix commonly derived from a Gaussian distribution.

## *3.4 Convolutional Neural Network (CNN)*

Convolutional neural network (CNN) is a deep learning algorithm whose main purpose is to process the structure of an array of data such as images, and this model is inspired by the organization of the animal visual vertex [20]. CNN is designed such that it can automatically learn spatial hierarchal features, from low to high hierarchies of features [22]. The main reason for opting for the CNN model in lung prediction is that it does not require hand-craft features extraction as done in various conventional machine learning classifiers such as random forest and support vector machine (SVM). The architecture of CNN is designed such that it does not require the segmentation of human organs by human experts. Finally, it is computationally very expensive as it has a lot of learnable parameters to estimate.

### 3.4.1 Architecture of CNN

The architecture of CNN for predicting lung cancer consists of four layers (Fig. 2).

1. *Convolution Layer:* This is the most significant portion of the CNN since it conducts feature abstraction by applying a tiny array of integers called a kernel to the input, which is a tensor [22]. At each tensor location, a product between each kernel element on an element-by-element basis and the output value is calculated by measuring and summing the input tensor, called a feature map, in the output tensor's corresponding location. Adding several kernels to an indefinite number of feature maps reflecting different input tensor properties replicates this process, allowing different kernels to be read as distinct feature extractors. Two

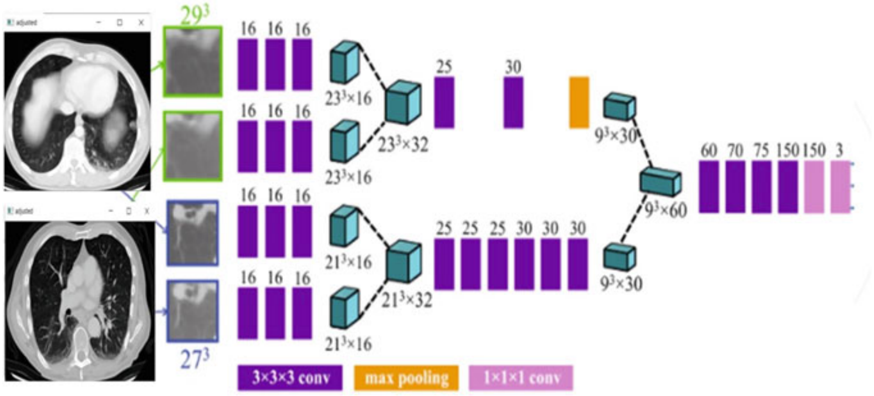


**Fig. 2** Convolutional neural network model

major hyperparameters that characterize the convolution process are kernel size and number. Usually, the former is 3 * 3, but often 5 * 5 or 7 * 7. The latter is arbitrary, deciding the depth of maps of the output function.

The key aspect of a convolution operation is weight sharing: Kernels are shared across all picture locations. Weight sharing produces the following properties in convolution processes: (1) Kernels travel across all image places and detect learnt local patterns, making local feature patterns collected by translation of kernels b invariant; (2) when kernels travel across all picture places and detect learnt local patterns, making local feature patterns collected by translation of kernels b invariant; and (3) model performance is improved by limiting the amount of parameters to be learned when compared to fully connected neural networks. The approach of training a CNN model for the convolution layer is to determine the kernels that work best for a certain goal based on a given training dataset. Kernels are the sole parameters in the convolution layer that are automatically learnt during the training phase. Hyperparameters like as kernel size, number of kernels, padding, and stride, on the other hand, must be specified before the training process begins.

2. *Activation Layer*: This matrix produced by the convolution layer generates a smaller matrix than the original one. Then, this matrix is passed through the activation layer which adds non-linearity into CNN that allows the model to train model through backpropagation. The activation function is represented by the rectified linear unit (ReLU) function.

$$f(x) = \max(0, x)$$

3. *Pooling:* To minimize the number of learnable parameters by adding translation invariance to small shifts and distortions, a pooling layer performs a conventional downsampling approach to reduce the function maps' in-plane dimensionality [22]. There is no learnable parameter in any of the pooling layers, but filter size,

phase, and padding are hyperparameters in pooling processes akin to convolution operations. The max-pooling strategy has been added into our model. It selects patches from the input function's maps, produces the patch's maximum value, and removes all other values. Maximum pooling using a filter of size 2 * 2 with a step of 2 is extensively used in practice. It selects patches from the input function's maps, produces the patch's maximum value, and ignores the rest. This reduces the in-plane dimension of the function mappings by a factor of two. In contrast to height and breadth, the depth dimension of function maps remains constant.

4. *Fully-Connected Layer:* The last convolution or pooling layer's output function maps are usually flattened, that is, turned into a one-dimensional (1D) number (or vector) array, and linked to one or more totally linked layers, also known as thick layers, in which each input is linked to each output by a learnable weight [22]. Once the features retrieved by the convolution layers and downsampled by the pooling layers are established, the network's final outputs, such as the probabilities for each class in classification tasks, are mapped by a subset of entirely connected layers. The number of output nodes in the final fully linked layer is usually equal to the number of groups. Each entirely linked sheet is followed by a nonlinear function, such as ReLU, as discussed previously.

### *3.5 Flask App*

Flask is a Python micro-framework that is commonly used to deploy ML models on the Web.[8] The model is trained, and validated images are passed on to the flask app to predict the type and severity of cancer which leads to real-time monitoring as any user can upload images and can get the results. For achieving the output, we have created two templates which include uploading the validated images and predicting the result (Fig. 3).

## 4 Results

### *4.1 To Improve the Quality of CT Scans Images Depicting Different Types of Lung Cancer*

In this study, we have worked on a simple and efficient method to improve the quality of the CT scans of lung cancer and the normal lungs. The main reason to improve the image quality was to achieve higher accuracy. The accuracy achieved by the CNN model before image improvisation was about 80% (refer to Table 1), but with the implementation of the image restoration technique, we achieved an accuracy of 90%

---

[8] https://www.kdnuggets.com/2019/12/excelr-deployment-machine-learning-flask.html.

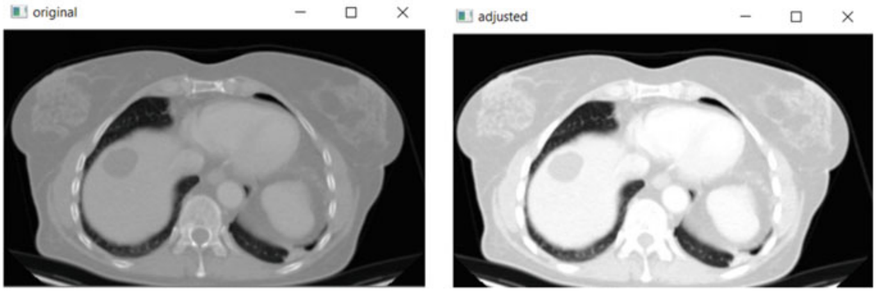
**Fig. 3** Snippet of flask app

**Table 1** Accuracy of the model before image improvisation

| Class | Split | Accuracy |
|---|---|---|
| 1. Adenocarcinoma | 320 + 80(400) | 256/320 (80%) |
| 2. Large cell carcinoma | 320 + 80(400) | 262/320 (82%) |
| 3. Squamous cell carcinoma | 320 + 80(400) | 260/320 (81%) |
| 4. Healthy lungs | 320 + 80(400) | 262/320 (82%) |

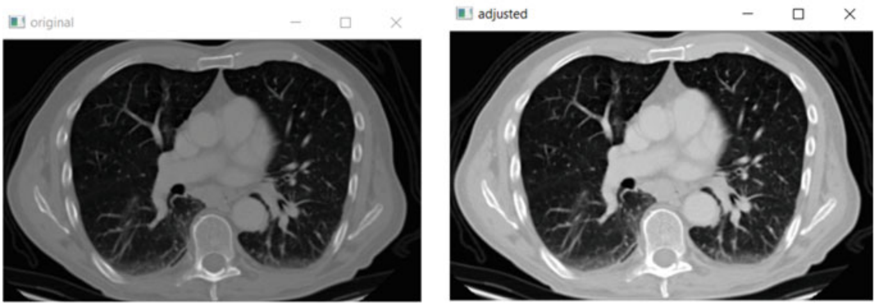
as mentioned below. Images are improved using the Weiner filter. The improved image quality for the different types of classes is represented in Figs. 4, 5, 6 and 7.

- *Adenocarcinoma:* Fig. 4 represents the improvisation of the image of adenocarcinoma lung cancer.



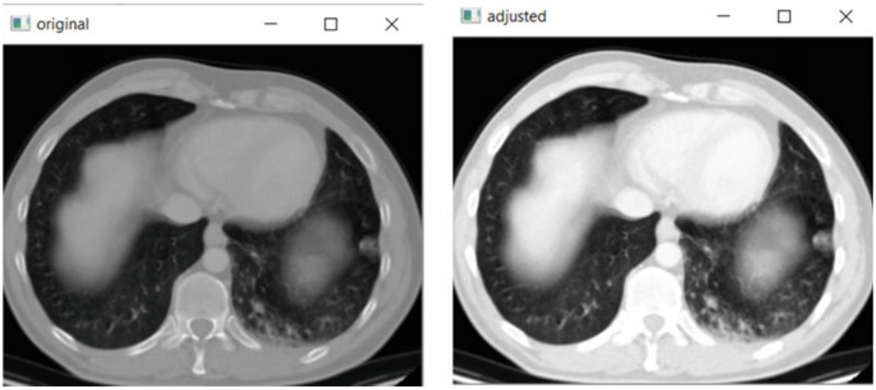
(a) (b)

**Fig. 4** Original **(a)** and improved **(b)** image of adenocarcinoma lung cancer



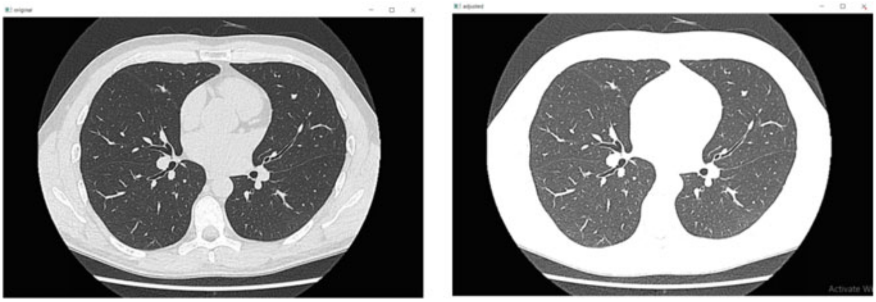
(a) (b)

**Fig. 5** Original **(a)** and improved **(b)** image of large cell carcinoma



(a) (b)

**Fig. 6** Original **(a)** and improved **(b)** image of squamous cell carcinoma



(a) (b)

**Fig. 7** Original **(a)** and improved **(b)** image of healthy lungs

- *Large Cell Carcinoma:* Fig. 5 represents the improvisation of the image of large cell carcinoma lung cancer.

- *Squamous Cell Carcinoma:* Fig. 6 provides the improvisation of the image of squamous cell carcinoma lung cancer.

- *Healthy Lungs:* Fig. 7 provides the improvisation of the image of healthy lungs.

### 4.2 To Classify the Lungs of a Person that is Infected with a Type of Lung Cancer and Severity Using a Two-Dimensional Convolution Neural Network (CNN)

To classify the lung cancer type, the improved images are feed into the 2D CNN model. Further, predictions were made by running the trained set on the images of the test set. The following snippets show the prediction results for the first three CT scans (Fig. 8).

Further, the accuracy of the model is observed. Accuracy defines the proportion of time a classifier function can correctly predict a class. The accuracy can be mathematically represented by the following equation:

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

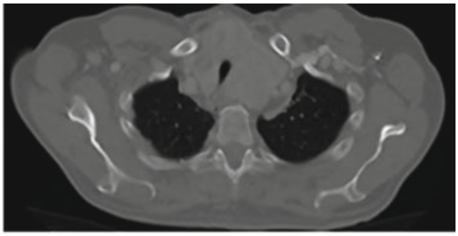
where true positive (TP) indicates the count of correctly detecting lung cancer, false positive (FP) is the count of correctly classifying healthy lung. True negative (TN) is incorrectly classifying lung cancer. False negative (FN) is incorrectly classifying the healthy lung. The accuracy achieved by all four classes, named adenocarcinoma, large cell carcinoma, squamous cell carcinoma, and healthy lungs, was above 90%. Further, Table 2 represents the accuracy achieved by our model concerning different classes present in our model.

Moreover, we have also calculated precision, $F$-measure, and recall to evaluate our model (refer to Table 3). Precision is described as a percentage of the retrieved occurrences that are relevant. It is the proportion of true positive cases predicted with lung cancer in our model. It can be defined mathematically as follow:

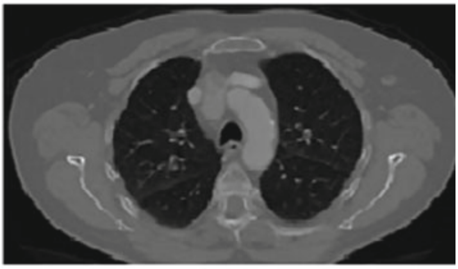
$$\text{Percision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

The recall is the percentage of the total number of relevant instances found. It is the ratio of real positive lung cancer cases predicted to the total number of positive cases in our model. It can be defined mathematically as follow:

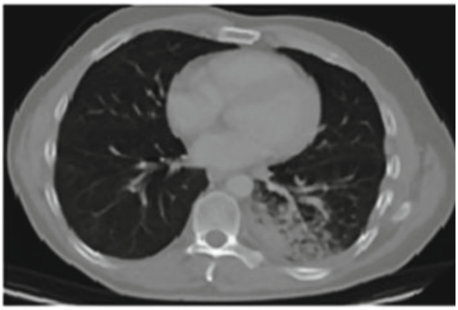
**Fig. 8** Prediction for test images



Large Cell Carcinoma



Adeno Carcinoma



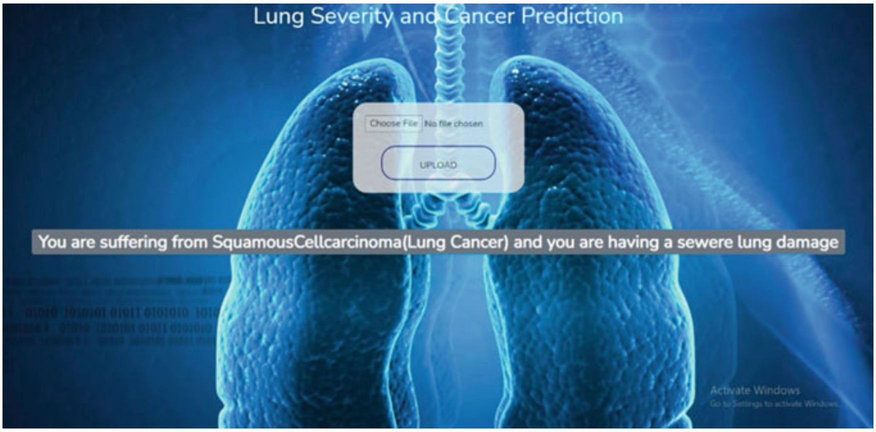
Squamous Cell Carcinoma

**Table 2** Lung cancer type and severity prediction using CNN

| Class | Split | Accuracy |
|---|---|---|
| 1. Adenocarcinoma | 320 + 80(400) | 288/320 (90%) |
| 2. Large cell carcinoma | 320 + 80(400) | 294/320 (92%) |
| 3. Squamous cell carcinoma | 320 + 80(400) | 291/320 (91%) |
| 4. Healthy lungs | 320 + 80(400) | 295/320 (92%) |

**Table 3** Classification performance measures for classifiers

| Classification measure | Performance |
|---|---|
| Precision | 0.99 |
| Recall | 0.90 |
| $F$-measure | 0.94 |


Lung Severity and Cancer Prediction

You are suffering from SquamousCellcarcinoma(Lung Cancer) and you are having a sewere lung damage

**Fig. 9** Prediction using flask app

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Precision and recall are achieved by *F*-measurement. It evaluates the test's precision and recalls while doing so. The *F*-score will provide a more realistic calculation of the success of the test by using both precision and recall. Mathematically, it can be described as follows:

$$F\text{ - measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Moreover, our research provides real-time monitoring due to the usage of the flask app that allows users to upload the images and get results (Fig. 9), whereas other studies are mostly static as no interaction with users is available. It is an effective method as saves a lot of time for doctors and can represent results quickly.

## 5 Discussion

In this work, we have collected CT scans of different lung cancer types from google images and also used a dataset from Kaggle. Further, image restoration is conducted using the Weiner filter technique to increase image quality. To predict the lung cancer type and severity using 2D CNN in which feature extraction is done automatically. Further, data augmentation can also be achieved using CNN which is not possible in other models. The accuracy achieved by our model is above 90%. Our research work is real-time monitoring as we have used the flask app to predict the lung cancer type by uploading the validated images.

In the previous works, a study in which features were extracted manually for the classification of lung cancer [4]. In comparison, in our work, we have inculcated a 2D CNN model that can automatically perform the feature extraction work. Whereas another study used three various supervised models such as linear regression, support vector machine, and artificial neural network to classify tumor biopsies as malignant or benign and achieved the highest accuracy with SVM [8], but the main drawback of this study is that feature extraction is done manually that is overcome by our study. Further, our study is monitored in real-time due to the usage of the flask app.

Moreover, one of the studies proposed a 3D CNN model to classify lung cancer but the accuracy achieved by this model is 86% which is very low as compared to our model [14]. Our model also detects the severity and predicts the type of lung cancer. Further, real-time monitoring is the main advantage of our model as it allows users to predict the cancer type with the usage of the flask app through the usage of CT scan images.

## 6 Conclusion

Cancers are one of the deadliest diseases as many people are dead worldwide every year. Lung cancer is a leading cause of mortality in men and women alike. The main goal of our research is to predict the lung cancer type and severity using a 2D convolutional neural network. The accuracy achieved by our model is above 90%. One unique feature is that our research is real-time monitoring with the usage of the flask app as it allows users to upload the CT scans of the lungs can predict the type of lung cancer and its severity. This study can help the radiologist to get immediate results about the type of cancer and its severity. The main limitation of this study is that only CNN is being used for the prediction, there might better other algorithms that can provide better accuracy. In future, we would like to evaluate other algorithms if they outperform the results of our model.

## References

1. Van Ginneken B, Setio AA, Jacobs C, Ciompi F (2015) Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In: 2015 IEEE 12th international symposium on biomedical imaging (ISBI). IEEE, pp 286–289
2. Barker J, Hoogi A, Depeursinge A, Rubin DL (2016) Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. Med Image Anal 30:60–71
3. Ojansivu V, Linder N, Rahtu E, Pietikäinen M, Lundin M, Joensuu H, Lundin J (2013) Automated classification of breast cancer morphology in histopathological images. Diagn Pathol 8(1):1–4. BioMed Central
4. Zubi ZS, Saad RA (2014) Improves treatment programs of lung cancer using data mining techniques. J Softw Eng Appl

5. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI (2015) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 13:8–17
6. Camarlinghi N, Gori I, Retico A, Bellotti R, Bosco P, Cerello P, Fantacci ME et al (2012) Combination of computer-aided detection algorithms for automatic lung nodule identification. Int J Comput Assist Radiol Surg 7(3):455–464
7. Choi WJ, Choi TS (2014) Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor. Comput Methods Programs Biomed 113(1):37–54
8. Sidey-Gibbons JA, Sidey-Gibbons CJ (2019) Machine learning in medicine: a practical introduction. BMC Med Res Methodol 19(1):1–18
9. Monkam P, Qi S, Ma H, Gao W, Yao Y, Qian W (2019) Detection and classification of pulmonary nodules using convolutional neural networks: a survey. IEEE Access 7:78075–78091
10. Paing MP, Hamamoto K, Tungjitkusolmun S, Pintavirooj C (2019) Automatic detection and staging of lung tumors using locational features and double-staged classifications. Appl Sci 9(11):2329
11. Pradeep KR, Naveen NC (2018) Lung cancer survivability prediction based on performance using classification techniques of support vector machines, C4. 5 and Naive Bayes algorithms for healthcare analytics. Procedia Comput Sci 132:412–420
12. Kirienko M, Sollini M, Silvestri G, Mognetti S, Voulaz E, Antunovic L, Chiti A et al. (2018) Convolutional neural networks promising in lung cancer T-parameter assessment on baseline FDG-PET/CT. Contrast Media Mol Imaging
13. Shiri I, Geramifar P, Bitarafan Rajabi A, Mohammadzadeh A, Pouraliakbar H (2018) Deep convolutional neural network (D-CNN) for efficient and automatic lung cancer detection. Insights Imaging 9
14. Alakwaa W, Nassef M, Badr A (2017) Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). Lung Cancer 8(8):409
15. Chaunzwa TL, Hosny A, Xu Y, Shafer A, Diao N, Lanuti M, Aerts HJ et al (2021) Deep learning classification of lung cancer histology using CT images. Sci Rep 11(1):1–12
16. Kaur J, Kaur R (2013) Biomedical images denoising using symlet wavelet with wiener filter. IJERA 3(3):548–550
17. Umapathi VJ, Narayanan VS (2014) Medical image denoising based on Gaussian filter and DWT SWT based enhancement technique. Int J Soft Comput Artif Intell 2(2)
18. Bala A, Hati C, Punith CH (2014) Image denoising method using curvelet transform and wiener filter. Int J Adv Res Electr Electron Instrum Eng 3(1)
19. Jadwaa S (2018) Wiener filter based medical image de-noising. Int J Sci Eng Appl 7:318–323
20. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48
21. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol 195(1):215–243
22. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629

# A Comprehensive Review of Brain Tumor Detection and Segmentation Techniques

**Ankush Azade, Praful Kumar, and S. Sowmya Kamath**

**Abstract** Brain tumors are particularly dangerous type of tumor, and if this is not treated in time it maybe prove to be deadly and may also spread across other body parts. Brain tumor is the swelling or growth of unwanted tissues in the brain that results from the unregulated and disordered division of cells. The presence of these tissues resulting abnormal behavior and lot of other complications. The detection of brain tumor is done by using different techniques out of which through magnetic resonance images (MRIs). The scanning process is a time-consuming manual task that needs the involvement of medical professionals. Automating the task of detection of the brain tumor while also grading the severity accurately can help in managing the patients' disease effectively. As tumor tissue of different patients is different, automating such processes is often a challenging task. Researchers have incorporated image segmentation for extraction of suspicious regions from MRI, using image processing and AI-based techniques. Radiomic analysis also plays a big role in feature extraction processes. In this paper, we present a comprehensive review of existing approaches for brain tumour detection, covering deep neural models, radiomic analysis and segmentation-based methods for brain tumor classification and segmentation, along with a discussion on prevalent issues, challenges, and future directions of research.

**Keywords** Tumour segmentation · Radiomic analysis · Medical image classification · Medical informatics

---

A. Azade (✉) · P. Kumar · S. Sowmya Kamath
National Institute of Technology Karnataka, Surathkal 575025, India
e-mail: ankush.202it003@nitk.edu.in

P. Kumar
e-mail: praful.202it020@nitk.edu.in

S. Sowmya Kamath
e-mail: sowmyakamath@nitk.edu.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_33

# 1 Introduction

The brain is the hub of all nerve activity, making it one of the necessary and delicate organs in the body. The spinal cord and brain constitute the nervous system of a human body, and the body's vital operations, such as thought, speech, and movement, are controlled by the central nervous system. Each year, about 350,000 new brain tumour cases are recorded around the world, with a 36% survival rate. A brain tumour is an unregulated cell multiplication that leads in abnormal growth. Abnormality in the brain and spine growth can result in negative effects on how a human walks, speaks, and performs basic cognitive activities. Regular headaches, speech difficulties, fatigue, difficulty in balance, inexplicable nausea, and vomiting are all common signs of a brain tumour. These symptoms are many times overlooked at early stages and they prove to be very dangerous later when detected. These tumours may often be divided into primary tumours and secondary tumours. Primary brain tumors are further divided into two categories: low grade and high grade. When compared to high-grade tumors, low-grade tumors grow more slowly. Secondary brain tumors are malignant tumors that originate in another part of the body before spreading to the brain. The severity of the condition determines the grade of the brain tumour, which ranges from Grade I to Grade IV.

- *Grade 1:* The brain cells seem normal and will grow gradually and slowly. At this stage, the long-term survival is very likely.
- *Grade 2:* The cells will start looking slightly abnormal along with slow growth of tumor. At this stage, the spread of tumor to the nearby tissues starts and it may grow to a life-threatening grade.
- *Grade 3:* The growth of tumor in the nearby tissues starts actively and by the looks the cells will look abnormal.
- *Grade 4:* At this stage, the growth of tumor is very fast and it will spread to nearby cells very rapidly.

Neurological exams such as computer tomography (CT), MRI, and other tests such as angiogram, spinal tap, and biopsy are used to detect brain tumours. MRI detection of brain tumours is one of the most effective methods, and current diagnosis process mostly involves trained human experts. Often, grading tumours is a difficult and error-prone process, due to the minute features that need to be identified for accurate grading. Automated detection and grading of brain tumor have thus received significant research attention, as early diagnosis and accurate grading can be very beneficial in managing patient treatment and improving disease prognosis. Detection at a benign stage (lower grade) can be treated very effectively and can help in successful prevention of the tumour turning malignant. Additionally, automated segmentation of MRI images of brain tumours can help enhance diagnostic and treatment choices, particularly in circumstances when access to qualified specialists and radiologists is restricted.

A comprehensive assessment of current state-of-the-art efforts for brain tumour classification and segmentation is offered in this paper. Deep learning-based works are primarily for classification of tumour in various ways, i.e., if the tumour is present or not and if present, classifying the tumor as per the severity grade of the tumour. Certain features of the brain tumour like the radius of the tumour, density of the tumour which can be recognized with radiomic analysis can help in the detection of tumour to a great extent, and these properties can be used in combination with one another to obtain good results. Segmentation techniques are used in order to separate the tumor from other parts of the brain in the MR image so that regions of interest can be identified. The remainder of this paper is structured as follows : Sect. 2 details the various approaches in deep learning for brain tumour detection. In Sect. 3, radiomic analysis-based techniques which are mainly used for feature extraction are discussed and segmentation-based approaches are discussed in Sect. 4. Also, various standard datasets available for the research community for brain tumour detection research are discussed in Sect. 5. Open challenges and research directions are discussed in Sect. 6, followed by concluding remarks and references.

## 2 Deep Learning Approaches for Brain Tumour Detection

There has been a lot of research in deep learning-based brain tumor detection, which can be perceived as a classification problem. Several studies have considered brain tumor detection as binary classification problem where two classes are *tumor* and *no tumor*. A few have addressed this task as a multi-class classification where different grades of tumor are considered as different classes. We present a review of existing works that address both binary classification and multi-class classification in this section.

Amin et al. proposed that for brain tumour detection, a seven-layer deep learning model of 3 convolution layers, 3 ReLU layers, and 1 softmax layer was used to segment and classify MR images. They segmented the input image into patches of 4 * 65 * 65 pixels, determined the centre pixel label for each patch, and gave it to the deep neural network. The overall results of segmentation and classification were enhanced using this way of partitioning the image into patches. They also assessed eight distinct benchmarks, all of which are widely used in the identification of brain tumours. The results on these bechmarks, namely BRATS2012, BRATS2012 (synthetic images), BRATS2013, BRATS2013 (synthetic images), BRATS2014, BRATS2015, ISLES2015, and ISLES2017, were in the range 95.5–100% which shows that model also generalizes very well and can be further used for real-time usage. Fusion of certain handcrafted techniques was addressed by Saba et al. for the prediction of glioma categories like high-grade and low-grade gliomas. High-grade gliomas have a two-year survival rate when treated early.

The authors used deep learning approaches, particularly convolutional neural networks, to recognize the patterns. The suggested technique offers a substantial contribution by transforming the input images into one channel. An accuracy of around 99% was obtained for the BRATS2015, BRATS2016, and BRATS2017 benchmarks from BRATS, but there was no testing done with other datasets, which limits the scalability of this work.

For predicting the progression-free and overall survival of a newly diagnosed glioblastoma patient, Chaddad et al. [1] proposed multi-scale characteristics based on the Laplacian of Gaussian (LoG) filter. T1-weighted imaging (T1WI) and fluid attenuation inversion recovery (FLAIR) images of 40 GBM patients were used to extract characteristics that were manually categorized as active tumour, necrosis, or edoema. To extract multi-scale texture features, the LoG filter was applied to each of these regions of interest, and then univariate approaches including the Spearman rank coefficient, logrank test, and Kaplan–Meier estimator were employed to validate the features. Multivariate analysis, as well as the relationship between them (i.e., random forest classifier). The first three and seventh features were shown to have statistically significant correlations with PFS and OS, respectively. The three characteristics derived only from the active tumor area were associated with OS (where $p$ is less than 0.05), and the hazard ratios (HR) were found to be 2.9, 3, and 3.24, respectively. Using the random forests classifier to predict the combined characteristics of patients with PFS and OS in GBM patients showed that the AUC was 85.37% and 85.54%, respectively.

Dastmalchian et al. adopted feature selection phase during which features which were discovered to be strongly associated were picked using Spearman's rank coefficient. One of the traits from the correlated pairs was chosen to avoid redundancy. Furthermore, the authors used the Wilcoxon rank sum test on each of the tumour groups for the specified attributes in the preceding steps. The text characteristics were evaluated using ROC and AUC. As a result, 20 out of 39 texture features were chosen for further investigation. When used to MRF, texture analysis is a well-established and powerful method for inferring local and regional heterogeneity in tumours, and it can reveal significant differences between tumour types. This work demonstrated the use of texture analysis on MRF-derived quantitative maps in malignant brain tumours. Hashemzehi et al. evaluated their proposed model on the T1-weighted CE-MR images and obtained a precision of 94.49%. In this model, a CNN and a neural autoregressive distribution are merged. They used a three-step methodology for their research that involved categorization, feature exploitation, and density estimation. The NADE model was initially trained to produce appropriate joint distribution, which made feature extraction easier later on. Separate CNNs were used to learn features from the produced output and the raw picture, limiting the impact of noise and rotation. Although the proposed model's operation was extremely distinctive, it was not tested on numerous datasets or datasets with a larger number of photos.

## 3 Radiomic Analysis for Brain Tumour Detection

Radiomics has been defined as "a quantitative approach to medical imaging that uses cutting-edge mathematical analysis to improve the data already accessible to physicians." It has been reported to have significant impact on clinical decision making as it utilizes the inherent image information like pixel intensity and spatial distribution for determining dominant features. Radiomics also aim to convert medical images into high-dimensional data. There are CAD systems that detect if tumor/cancer is present or not, radiomics are directed towards extracting large number of features that may be used for different purposes. With the help of radiomics, decision support tools that may help in accurate prediction of tumor can be developed. Radiomic features can be combined along with other characteristics of the patients in order to improve decisions. As discussed the process of feature extraction includes six steps: acquiring images, identifying the interest areas in image, segmentation of volumes, extracting and qualifying descriptive features, using features to for collecting database, mining this data for the development of classifier models.

Brunese et al. employed ensemble learning to combine ten distinct machine learning algorithms, including closest neighbor, Linear SVM, RBF-SVM, neural network, Gaussian process, naive Bayes, C4.5, random forest, QDA, and logistic regression. First order, shape, the co-occurence matrix for gray levels, the run-length matrix for gray levels, and the gray level size zone matrix were also employed as radiomic features. The first step was to build a classifier using the previously mentioned machine learning techniques. The second step in this process was to choose the best classifier based on the criteria that were considered, such as decision boundaries, performance obtained by single models, and calibration plots. The voting ensemble classifier was used to predict the final prediction from among the many predictions provided by the best classifiers. Following that, they built a model in two stages: training and testing. In addition, the dataset was divided into three parts, namely training, testing, and validation, in the ratio 2:1:1. The accuracies obtained for Grade 1, Grade 2, Grade 3, and Grade 4 were 95.1 %, 95.1 %, 95.2 %, and 95.1 %, respectively, because they predicted four different grades of brain tumour.

Supervised learning was used by Mercaldo et al. for the prediction of the type of lung cancer among the five categories from the magnetic resonance images. A total of 30 radiomic features were considered which belong to two basic categories, namely shape and gray level size zones. Carefully considering the radiomic features is also very crucial stage as it will affect the final prediction to a very large extent. The neural network considered was a basic single layer feed forward perceptron with the Heaviside step function as the activation function. The precision obtained was maximum of 93.1% but since the prediction was to identify the different stages of lung cancer the study is very useful. Gillies et al. showed that features can be extracted for varied imaging modalities like position emission tomography, magnetic resonance, and computed tomography. Radiomics generally have six steps: collection of image, identification of volume of interest, volume segmentation, extraction of different features, using these features for database preparation, and development of classifiers based on these databases. The various usecases of radiomics include enabling diognosis, tumor prognosis, treatment selection, deciding on location of biopsy or resect. However, radiomics is a very young discipline and there are certain challenges faced like the reproducability of data, data sharing, lot of available data(big data), and lacking of standards. In the future, radiomics can prove to be a field of great use in medical health care.

## 4 Segmentation Approaches for Brain Tumour Detection

Segmentation is the process of assigning a specified set of classes to each pixel in an image and is an important tool in general medical image analysis tasks. It is predominantly used for extracting targeted information like regions of interest (RoI) from medical diagnostic scans. Various segmentation methods like watershed, threshold, edge-based, region-based, and clustering-based segmentation can be used. Deep neural networks have recently been widely employed for RoI segmentation as well. We discuss some relevant works that adopt these methods for brain tumor detection below.

Watershed segmentation is based on topological interpretation where resulting boundaries form closed and connected regions. Jemimma and Vetharaj [2] proposed an approach using watershed dynamic angle projection and CNNs, for accurate measurement of tumor regions. The textured characteristics of the brain were extracted using the DAP pattern, and the tumor and non-tumor areas of the MRI brain picture are classified using the CNN. Threshold segmentation is used to break down images into their constituent segments or areas in order to recognize objects and edges is known as image segmentation. The pixels in the image are partitioned based on their attributes. The pixels in the image are either replaced with black or white in threshold segmentation. Natarajan et al. [3] proposed a method based on sharpening and median filtering for preprocess and to improve image used histogram equalization and used thresholding for segmentation of image. Finally, using the image subtraction approach, the tumor area is derived. Threshold segmentation is not applicable when the object area ratio changes.

Edge-based segmentation is based on the detection of discontinuities. Maiti et al. converted an RGB color image to an HSV color image, it has three sections: color, saturation, and intensity. After enhancing the contrast, the watershed algorithm is applied to each section of the image. The disadvantage of this method is that it is difficult to construct a robust edge linking. An picture is divided into homogenous areas using region-based segmentation techniques. In this strategy, unique object boundaries are generated by flow from an interior point to an outside region. Tak et al. [4] proposed object labeling for more specific information about the tumor site and for detection and extraction they used MATLAB software. SVM was utilized in an unsupervised way to develop and retain the pattern for future usage.

Clustering-based segmentation typically employs unsupervised approaches for segmenting RoIs. Maksoud et al. [5] proposed well-formed hybrid segmentation techniques, based on fuzzy $C$-Means algorithm. They found that this did not correctly segment images that are influenced by noise, outliers, or other imaging distortions. They included a 3D evaluation of brain tumor identification using 3D slices. To recognize edges in CT and MRI images, Sharma et al. [6] used the Gabor transform in association with soft and hard clustering. Their study included a total of 4500 MRI scans and 3000 CT pictures.

$K$-means clustering was used to divide related subgroups of characteristics. The author used the Fuzzy $C$-Means algorithm to represent the images as histogram properties. Islam and Ahmed [7] proposed a comparison of $k$-means, $k$-medoids, and hierarchical clustering algorithms. They reported that the $k$-means clustering method outperformed other methods and was simple to use. Bhoi and Gupta [8] for investigating two clustering segmentation techniques, $k$-means and fuzzy $c$-means, in terms of execution duration and computation of the tumour area. Because it takes less time to execute and extracts the tumour region more precisely, the $k$-means segmentation method outperformed fuzzy $c$-means. One of the most effective image classification strategies was , which was used to create a fused segmentation. The ensemble methodology, which is based on the potential field notion of physics, combines data from PFS and other methodologies. The authors reported an accuracy of 89.8% using the BRATS dataset for experimental validation of their approach.

Deep learning-based segmentation techniques have been adopted by several researchers in recent years. Ahuja et al. [9] applied transfer learning for brain tumor identification and super-pixel approach for segmentation. They created a model based on the BRATS 2019 brain tumor segmentation competition. For training purposes, they used VGG19 transfer learning and with the help of super-pixel method tumor was separated into LGG and HGG images. Jia and Chen [10] used support vector machine to construct a heterogeneous segmentation that is completely automated. To train and test the accuracy of tumor diagnosis in MRI images, they used a classification approach known as probabilistic neural network classification system. This model focused on automatic meningioma segmentation and employed a multispectral brain dataset. Mittal et al. introduced the notion of the stationary wavelet transform (SWT) and the novel growing convolution neural network (GCNN). The emphasis was on enhancing the convention system's correctness. In addition, a comparison study was conducted using SVM and CNN, which revealed that the suggested model outperformed both in terms of accuracy. Convolutional neural networks are compared to a combination of texture-based feature extraction methodologies and conventional techniques by Affoso et al., using different learning models for the classification of biological images. This study shows that deep learning-based algorithms like CNN are the best at image classification.

## 5 Datasets

Several data sources are currently available, that provide images from brain tumor of varying tumor ranges. The most commonly used modalities are magnetic resonance images (MRIs). Some of the most common resources are BraTS, REMBRANDT, images from radiopaedia, etc. RAMBRANDT and images from radiopaedia repository contain images that are divided into 4 grades of tumors. The RAMBRANDT dataset contains 110,020 MR Images with Region of Interest segmentation and medical reports from 130 different patients. The 4 categories in which the dataset is divided are I, II, III, and IV depending on the grade of the tumor and they contain MR images from 37, 40, 25, 28 patients, respectively. The radiopedia repository dataset contains images from 121 MRIs which are labeled with four labels divided into I, II, III, IV categories and they contain 36, 32, 25, and 28 MRIs, respectively, for the given grades.

BraTS 2019 [11] makes use of data from multiple institutes, the data is preoperative MR images, and the main focus is on instrinsically heterogeneous brain tumors, specifically gliomas. Furthermore, BraTS’19 focuses on predicting patient overall survival using integrative analyses of radiomic features and machine learning algorithms in order to pinpoint the clinical relevance of this segmentation task. Finally, at BraTS’19, an experimental evaluation of tumour segmentation uncertainty will be carried out. The dataset is divided into four subfiles, with dataset volumes classified as T1, T2, Flair, or T1Ce sequences. There are 155 slices in each volume. MRI images are used to create the slices. One volume necessitates 155 slices. Glioma dataset is with 210 high-grade volumes. There are 285 volumes in total, with 75 volumes in another type of glioma. Because it is a standard dataset for detecting brain tumours, BraTS2019 can be used. The Kaggle dataset [12] is containing MRI images with 155 tumor and 98 no tumor images (253 total images) (Figs. 1 and 2).

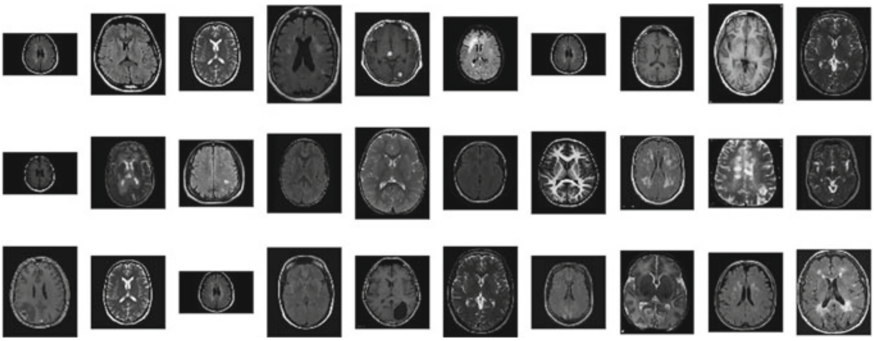


**Fig. 1** Healthy brain MRI images

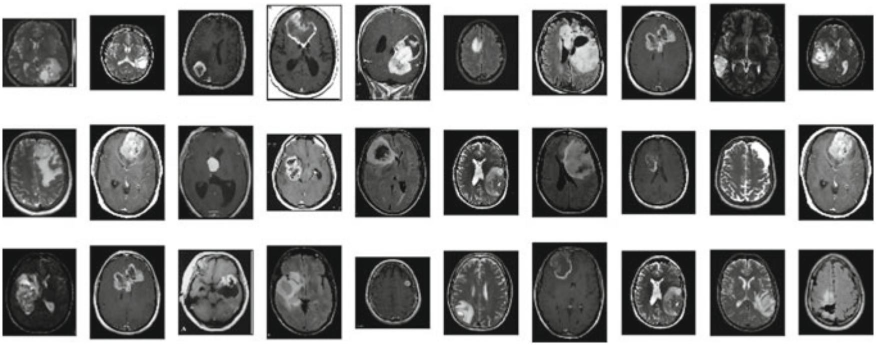


**Fig. 2** MRI images containing tumor

## 6 Open Challenges and Research Directions

Based on the detailed review of various approaches for brain tumor detection, several interesting insights were revealed. It was observed that deep learning techniques have been extensively used for accurate classification of brain tumor into multiple classes, for enabling automated severity grading tasks. This is critical in clinical workflows for supporting decision-making tasks that can help improve the diagnostic process and reduce time taken for diagnosis. Most studies have focused on binary classification of the brain tumor that reduces the task to a prediction with respect to if brain tumor is present or not. However, the main concern is the type of brain tumor as there is a very high chance of grade 1 and grade 2 tumors not being detected in time. Also, for many of the datasets the number of samples in different classes is very unbalanced. Although the classes can be balanced using augmentation techniques, the sample images obtained after augmentation are just copies of the other images by introducing some small variations. However, if there can be datasets with almost balanced classes, then it would make the results better and more practical to use in real-world applications (Figs. 3 and 4).

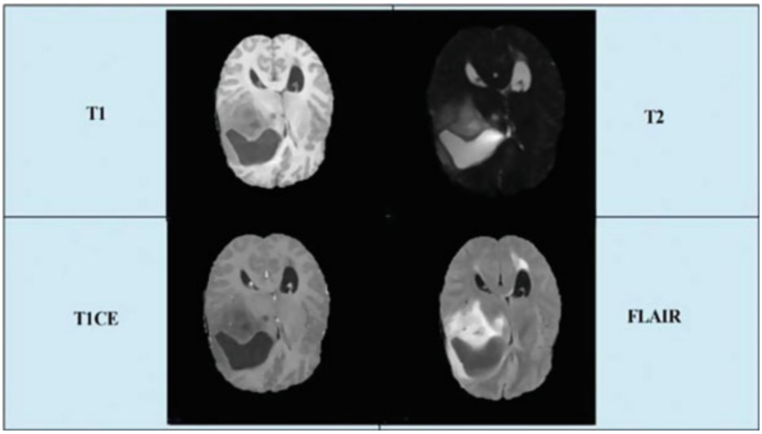


**Fig. 3** Sample images from BraTS dataset

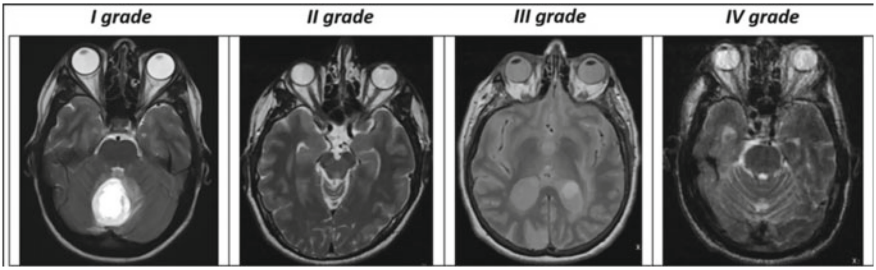


**Fig. 4** Images of 4 grades of brain tumor

Segmentation is a technique in which an image is divided into number of partitions called as segments. The image segments simplify image processing and analysis by reducing the complexity of the image. Most existing studies have focused on standard parametric methods like threshold, watershed, region-based, and clustering-based segmentation methods. Threshold segmentation is an easy-to-use method which takes pixels' mean value as threshold value, the disadvantage being that it is not suitable when the object area ratio changes. Watershed is based on topological interpretation which results in closed and connected regions and the problem is it gives excessive segmentation while edge-based employs discontinuity detection and performs better in photos with strong object-background contrast. In these standard techniques, $k$-means clustering is easy to implement, takes less processing time, and also provides accurate predictions. While applying these segmentation techniques on the MRI images, skull stripping has not been used. This is a significant gap as image quality has a significant impact on the segmentation method's performance. The result of image segmentation is determined by a lot of aspects, including pixel color, intensity, image similarity, and issue domain. Hence, automated methods that can address these issues are the need of the day.

## 7 Concluding Remarks

There are different methods for identifying brain tumor by looking at magnetic resonance images. Deep learning methods and segmentation along with radiomic analysis play a huge role in the detection of brain tumor. Deep learning methods majorly focus on learning of a model by learning from different images given as a train data. Segmentation focuses on division of regions in an image so that it can identify the tumor by those regions. Radiomic analysis is the extraction of features like shape and size of tumor from a given image so that these features can be used for the further training and testing of data. These methods can be collectively used to obtain good results. Many of the publications combined deep learning with radiomic analysis, allowing characteristics derived from images to be input into the model for training and testing. These give better results as compared to the studies that have individually used one of these methods.

## References

1. Chaddad A, Sabri S, Niazi T, Abdulkarim B (2018) Prediction of survival with multi-scale radiomic analysis in glioblastoma patients. Med Biol Eng Comput 56(12):2287–2300
2. Jemimma TA, Vetharaj YJ (2018) Watershed algorithm based DAPP features for brain tumor segmentation and classification. In: 2018 International conference on smart systems and inventive technology, pp 155–158
3. Natarajan P, Krishnan N, Kenkre NS, Nancy S, Singh BP () Tumor detection using threshold operation in MRI brain images. In: 2012 IEEE International conference on computational

intelligence and computing research
4. Tak R, Kumar N, Satyaki, Verma S, Dixit S (2017) Segmentation of medical image using region based statistical model. In: 2017 International conference on intelligent computing and control systems, pp 1247–1253
5. Maksoud E, Elmogy M, Al-Awadi R (2015) Brain tumor segmentation based on a hybrid clustering technique. Egypt Inf J 16(1):71–81
6. Sungheetha DA (2020) Gtikf-gabor-transform incorporated k-means and fuzzy c means clustering for edge detection in ct and mri. J Soft Comput Paradigm (JSCP) 02:111–119
7. Islam S, Ahmed DM (2013) Implementation of image segmentation for natural images using clustering methods. Int J Emerg Technol Adv Eng 3
8. Krishnakumar S, Manivannan K (2020) Efective segmentation and classification of brain tumor using rough k means algorithm and multi kernel svm in mr images. J Ambient Intell Humanized Comput 6751–6760
9. Ahuja S, Panigrahi B, Gandhi T (2020) Transfer learning based brain tumor detection and segmentation using superpixel technique. In: 2020 International conference on contemporary computing and applications, IEEE
10. Jia Z, Chen D (2020) Brain tumor detection and classification of mri images using deep neural network. IEEE Access ( Early Access ) 1–1
11. BraTS (2019) https://www.med.upenn.edu/cbica/brats2019/data.html
12. Kaggle Dataset. https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection
13. Irsheidat S, Duwairi R (2020) Brain tumor detection using artificial convolutional neural networks. In: 2020 11th International conference on information and communication systems, IEEE
14. Rajan P, Sundar C (2019) Brain tumor detection and segmentation by intensity adjustment. J Med Syst 43
15. Elmegy MT, Magad KM-E (2020) A multiple classifiers system for automaticmultimodal brain tumor segmentation. In: 2019 15th International computer engineering conference, IEEE, Cairo, Egypt
16. Agrawal R, Sharma M, Singh BK (2017) Segmentation of brain tumor based on clustering technique: performance analysis. J Intell Syst 28
17. Sreeji VR, Beevi A, NN (2013) Survey on different methods of image segmentation. Int J Sci Eng Res 4
18. Zaitoun N, Aqel M (2015) Survey on image segmentation techniques. Int Conf Commun Manage Inf Technol 65:797–806
19. Chakrabarty N Brain mri images for brain tumor detection. https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection

# Early Depression Detection Using Textual Cues from Social Data: A Research Agenda

**Sumit Dalal, Sarika Jain, and Mayank Dave**

**Abstract** Mental disorders are one of the most prevalent diseases in the modern world, forcing people to live a life of disability, and hundreds of work hours are wasted. Late detection of a disorder makes it become a severe problem. Early detection can help prevent the problem, and if prevention is not possible, at least intervention could be taken to avoid a serious problem. The recent approaches of using social media posts of people and machine learning or deep learning for mental assessment have given hope to fight the situation. From the literature, we found the problem is sufficient data for applying these approaches, especially in the mental health domain. Another issue is the datasets imbalanced number of diagnosed and control users. For the issues, we experimented with taking social data for varying numbers of users and then differing numbers of posts for them. Also, the number of diagnosed and control users is taken in different ratios. We found that the accuracy metric of the results is impacted though it is not that much. We have also proposed to use peephole BiLSTM and presented its results.

**Keywords** Mental health · Neural network · Depression · Word embedding · Machine learning · Psycholinguistics

## 1 Introduction

Mental health is a problem targeted by the reach of the Internet. A 2001 report by World Health Organization (WHO) says that approximately 450 million people were having mental problems at that time [1], and depressive disorders would be the

---

S. Dalal (✉) · S. Jain · M. Dave
National Institute of Technology Kurukshetra, Kurukshetra, Haryana 136119, India
e-mail: sumitdalal9050@gmail.com

S. Jain
e-mail: jasarika@nitkkr.ac.in

M. Dave
e-mail: mdave@nitkkr.ac.in

S. Jain et al. (eds.), *Proceedings of the International Health Informatics Conference*,
Lecture Notes in Electrical Engineering 990,
https://doi.org/10.1007/978-981-19-9090-8_34

second-highest disease by 2020. Bipolar disorder, dementia, and depression are the most prevalent disorders with approximately 60 million, 48 million, and 21 million patients. Depression is one of the 20 main reasons leading people to suicide and accounts for approximately 1.5% of suicides in the world [2]. Coronavirus disease 2019 (COVID-19) has exacerbated the situation. Extended lockdowns have made this situation more critical. According to a survey conducted by US Census Bureau in December 2020, 42% of the surveyed adults in the US have shown symptoms of depression or anxiety compared to 11% in January–June 2019 [3]. World Bank has included mental health in the Sustainable Development Goals list.

The mental health assessment is very clinic centric. But a person suffering from depression generally avoids visiting a psychiatrist for various reasons (cost, social stigma, etc.). In previous studies [4–6], we have examined the works assessing mental health from various inputs. The researchers found that sensor data, social data (text, images), and the Internet search history of a user can be utilized for the mental health assessment of the user [7]. Authors also used search queries' data of Internet users for their mental health assessment [8, 9]. As sensor data (smartphone, smartwatch, or wristband) is not accessible easily; researchers focused publicly available user-generated content on social networking sites (SNSs) (Twitter, Reddit, etc.) to detect and monitor a mental disorder. The employment of user-generated data in the mental health analysis process assists in manipulating issues of the traditional mental healthcare system. Authors used various forms of social data like images [10], music [11]. As most users share activities via writing a post, status, and comment, textual data is the widely explored form in the domain contrasted to other forms (emoticons, audios, and videos). Linguistic features and affective user profiles have been extracted from the textual social data and used in the process.

The data generated on social platforms is of big data form. Statistical approaches are not much popular to analyze social data due manual intervention and time-consuming factors. This leads a transition to machine learning (ML) or deep learning (DL) methods. Lexicon-based approaches have been applied frequently in combination with machine learning algorithms. Some previous work builds multiple lexicons (depressants, negative emotion, etc.) and considers terms from these as features in the classification task. Automated analysis of text related to mental health aims to help medical decision-making, public health policies, and improve health care. Early detection is an essential aspect of mental assessment from social data and is paid much attention to while implementing ML and DL algorithms.

**Our Contributions:**

- Investigation of data requirements for depression classification based on the number of users is taken out.
- We also examined the impact of the number of posts per user for different machine learning and deep learning methods.
- An agenda for future research aspects is presented.

The paper is organized into six sections, including Introduction. Section labelsec2 presents fundamental terminologies of the domain. Section 3 discusses the depression classification related works. In Sect. 4, dataset building process and preprocessing of the same are given. Section 5 provides data requirements investigation and the results achieved. Conclusion of the study is presented in Sect. 7. In Sect. 6, a research agenda is given out.

## 2 Technical Background

The section introduces technical aspects of mental assessment from social data with machine learning or deep learning. A small introduction to terms like dataset, preprocessing, feature extraction, etc., is presented here.

### *2.1 Dataset*

People explored the use of social media data in mental health assessment. Social posts language is a lens to user's cognitive behavior and does not interfere with their daily activities. Social data can be text, images, emoticons, and videos. We employ two datasets from Twitter and Reddit. We have examine the effect of balanced and imbalanced datasets on machine learning and deep learning models. When the numbers of positive and negative cases are approximately the same, the dataset is balanced. But if the ratio of positive to negative cases is more or vice-versa, the dataset is imbalanced. An imbalanced dataset made the machine biased toward a class, which enlarged negative true and false positive.

### *2.2 Feature Engineering*

Feature engineering involves tokenization, stop word, and unnecessary words (hashtags, URL, foreign words, and cardinals), removal, stemming, lemmatization, negation handling, and sarcasm handling. Stop or unnecessary word removal, stemming, and lemmatization help in reducing the dimensionality of feature vector space. However, it should be kept in mind that stop word removal may invert the sense of a sentence. It may be helpful only in some cases where two negations got canceled. So negations should be handled carefully. Sarcasm is irony or implicit negation and is challenging to find out if a sentence is sarcastic. For this purpose, knowledge of user behavior and sentence context could be used.

## 2.3 Feature Encoding

Mainly, *one-hot encoding, bag of words*, or *Tfidf* (or some Tfidf alteration) are employed to transform the natural language features into machine acceptable form. One-hot encoding only informs about the presence of a feature/word in the record, whereas bag of words (BOW) or term frequency-inverse document frequency (Tfidf) tells frequency of a feature, which is important in sentiment analysis or mental health classification. Tfidf gives higher importance to discriminative features/words that appear in specific documents.

The *word embedding* layer in neural networks provides an extra sheet for models to understand the data. The embedding layer can be initialized with all zeros or the pre-trained word embedding vectors (Word2Vec, GloVe, and fastText). Word vectors are generated as a byproduct of classification or some other task through a shallow neural network. A word is placed in space at some position with respect to other words. With respect to whom a word is placed, the number of words is decided as per the size of the dataset or task. In simple language, embedded vectors represent a word in a fixed number of dimensions, and each dimension is a word. Value for a dimension is decided based on the semantic relationship of dimension and the word to be represented. Two words with similar meanings are assigned similar values for all dimensions, while words with opposite or indifferent meanings are placed far in the dimension space.

## 2.4 Approaches

Analyzing big social media data manually is not feasible. Researchers consider methods like lexicon-based, machine learning, and deep learning. In ***lexicon-based*** methods, dictionary of particular terms (like depressants or negative emotion words) is generated to check user posts for their presence. From lexicons, researchers transitioned to ***machine learning*** (ML), where support vector machine (SVM) and Naïve Bayes (NB) are popular algorithms. A dataset with positive (people diagnosed with mental illness) and negative cases (people without mental illness) is used to make the machine learn about the features and differentiate the posts based on the learned features. We have considered SVM and NB in the baseline category.

From ***deep learning*** architectures like simple neural networks (NNs), convolution neural networks (CNNs), and recurrent neural networks (RNNs), RNN performs better when it comes to contextual and sequential information processing in natural language processing (NLP) tasks like machine translation, sentiment analysis, etc. Simple RNN architecture reminds sequential information for small time steps. While peephole and GRU variants of *long short-term memory* (LSTM) remember information for long enough. *Bidirectional networks* forward and backward units are placed together to consider future time steps. These networks assign same weight to all hidden states while producing context vector which is not good. Hidden states (words) can be assigned weights depending on their relevance in the sentence. ***Attention*** provides higher priority to important parts of a sentence, and these can be interpreted as reason for the result. ***Hierarchical network*** leverages different or same types of neural networks stacking them together. One such prevalent network is built by placing BiLSTM at one layer and attention at the other. Output from the first layer is input to the second layer

## 3 Related Work

Many authors chose neural networks for mental disorder classification. CNN with pre-trained GloVe is used to perform multi-class classification [12]. A convolutional neural network with encoders is implemented to generate attributes from posts, images, and interaction threads by social media users to detect stress [13, 14]. Authors in [15, 16] experimented with pre-trained and specifically trained fastText and GloVe word embedding for depression detection with CNN. Word vectors trained on Twitter dataset inputted to BiLSTM for depression classification [17].

Some studies have hierarchically placed more than one type of neural networks. A hierarchical architecture is presented in [18] to handle the imbalanced dataset. Positive posts from the XGBoost classifier are passed to attention-BiLSTM with an embedding layer to calculate post vectors by aggregating word-level importance vectors. While in [19], authors used bidirectional gated recurrent unit (GRU) as hierarchical recurrent neural network (HRNN) with attention layer on Reddit dataset to perform multi-class classification. Pre-trained word vectors are given as input to calculate aggregated sentence representations which are further combined to find document representations. A bidirectional gated recurrent unit (GRU) is employed in [20] to assign attention weights at word and sentence level as hierarchical attention network (HAN). The authors also performed word level attention weight analysis has been done for relevant phrase extraction. Hierarchical neural network with CNN and LSTM is employed to predict self-harm depression level [21].

In [22], the authors focused on various word clustering approaches, including Word2Vec and Glove, on the Twitter dataset to detect PTSD and depression. They considered unigrams, metadata, topics' features, and ensemble classifiers using different textual feature sets. Attention layer above Bi LSTM with pre-trained global vectors (GloVe) (for all words, positive emotion words, and negative emotion words) is used to get contextual and emotional semantic information of a post [23]. Authors in [24] used depression lexicon (depression-related unigrams) and UMLS concepts as input features to gated recurrent unit (GRU) for depression detection. The only feature vector used was bag of words and hence no embedding layer. For post-level classification of depressive tweets, authors in [25] applied a simple LSTM RNN on a balanced dataset while tuning hyper-parameters.

*Scope* Authors from simple neural networks to convolution neural networks (CNNs) and recurrent neural networks (RNNs) with and without word embeddings are manipulated by researchers. BiLSTM is the mainly focused RNN variant in the domain, but peephole BiLSTM is missed. Also, specifically trained word vectors are not focused much yet. The problem of data sufficiency and imbalance is mainly focused on in this work.

## 4 Preparing Data

The dataset is created from downloaded Twitter posts while following the process mentioned in [26]. Following sub-sections will discuss the process of dataset building, data preprocessing, and lexicon building.

### *4.1 Building Dataset*

While building the dataset, we have decided to get on SNScrapper due to limitations on the total number of posts downloadable from Tweepy SNScrapper allows streaming of more than 3200 posts per user available publicly. Stored tweets are analyzed manually to abandon posts with advertisements or fake mental health statements, and a group of genuinely appearing diagnosed users is collected. We cannot verify if their statements on Twitter are genuine or not. For authenticity, we depend on the user's self-reported diagnosis statements on Twitter, like "I was diagnosed with depression." Users mentioning diagnosis at a very early stage of life or many years ago are discarded to keep noise low. Posts in response to specific campaigns are also removed to limit fake self-diagnosis statements. We endeavor to create data with as much low noise as possible, but we cannot ignore the fact that we only depend on the user's self-diagnosis statement. All user posts are merged into one document and considered a single post. Other than posts, we have not considered other features like followers, following, likes, etc. We have considered three experiments. (1) The first experiment is to learn the ideal number of users for the data scarcity problem in the domain and (2) To find an optimal number of posts per user required for implementable physical world systems.

For the experiments, various versions of the dataset have been prepared. We took 150, 350, and 550 users from diagnosed and control groups into a one-to-one ratio for the first experiment. For the second experiment, the dataset versions contain 1000, 3000, and 5000 most recent posts for a user for approximately 550 depression diagnosed users and a one-to-one ratio of control users, i.e., balanced versions. One more condition post up to 2 years ago is downloaded in the dataset in all versions.

### *4.2 Preprocessing*

Though data preprocessing has a significant role in decreasing noise and the number of features (dimensionality reduction), which steer better accuracy and low time complexity of the system, we have done only simple filtration to keep sentimental and psycholinguistic information in posts intact. Regular expressions have been built to remove special symbols, numerical values, blank spaces, URLs, hashtags, retweets signs, and other non-informational (in our context) signs. Contractions library employed for converting words in original form from the contracted one like cannot to can not. Stop list of words nltk library has been modified; negative sentiment and personal pronoun words are not removed as previous research showed more usage of these features in the posts of depressed persons. WordNetLemmatizer from nltk text processing library in Python is considered for lemmatization. Lemmatization is better than stemming as it takes part of the speech (POS) parameter.

### *4.3 Lexicon Building*

Dictionary-based approaches were followed from the Naïve times of natural language processing even before using machine learning in the domain. Dictionaries employed contain words related to specific categories, and then people messages/posts are searched for those terms. The number of time's words from a particular lexicon found in user posts help in calculating the consumption frequency of the group user belongs to. To differentiate diagnosed users from the control one, we have to inspect the consumption of positive sentiment, negative sentiment, and negation words in posts. To accomplish this motive, we have formatted three lexicons (positive_sentiment_words, negative_sentiment_words, and negation_words) with a total of 440, 164, and 21 words, respectively, from the Internet (just the Internet by Google).

## 5 Examining the Data Needs of Various ML and DL Methods

We have done three distinct experiments on two aspects which are number of users and posts per user and analyzed results produced by machine learning and neural network methods. We chose different machine learning and deep learning learners like Naïve Bayes, SVM from machine learning and hierarchical attention network (HAN), simple BiLSTM, BiLSTM with Word2Vec, and proposed peephole BiLSTM attention network.

We have observed some facts regarding the language of depressed and control users from the preprocessed Twitter dataset. We find that control group users write longer posts compared to the diagnosed group users. On an average, the control users have 19,768 words as the average length of their posts while diagnosed group has 17,247. Number of posts considered for these statistics is 3000 posts per user. Frequency of other linguistics blocks also varies between the two groups like number of personal pronouns, positive emotion, negative emotion, and negation words are more for control group users than diagnosed group, but when compared based on the average length of posts in the groups, these averages are almost similar in both the groups.

Like, 2.13% positive emotion words out of average words are found in diagnosed group user's posts while 2.01% is what we cannot inference any specific result from this metadata analysis.

## *5.1 Benchmark Model Summary*

The literature found that support vector machine (SVM) and Naïve Bayes algorithms are implemented widely compared to other machine learning algorithms due to simplicity and better results. Besides these, we implemented algorithms from neural networks like simple neural network (NN), convolution neural network (CNN), BiLSTM with Word2Vec, and peephole BiLSTM with attention layer and Word2Vec.
*SVM NB* After preprocessing, we need to convert social posts from natural language to mathematical numbers as a machine does not understand natural language. For these algorithms, posts have been converted into vectors by Tfidf vectorizer from Sklearn.

*Simple NN* has 128, 64, and 1 neuron layers, respectively, with 256 epochs and 128 batch sizes. For text to sequence conversion, tokenizer from Keras is imported.

*CNN* Besides a one-dimensional convolution layer of 128 neurons and two dense layers of 10 and 1 neurons, CNN text contains one one-dimensional global max pooling, flatten, and embedding layer each. Here, values for batch size and epochs are 100 and 5.

*BiLSTM Networks* All variants of BiLSTM Networks are built with four layers: embedding, bidirectional LSTM, and two dense layers. Bi LSTM layer contains 64 neurons, while dense layers have 10 and 1 neurons. In BiLSTM variants, the batch size is 64 and the epochs size is 5.

Word vectors from pre-trained embeddings are input in all neural network variants except Simple NN. The dimension of vectors sets to 50 for all the models with embedding layer. All posts of a user are taken as a single document for classification. The maximum document length is 20,000 due to system and time constraints. Even though 73,124 and 19,768 are maximum and average words in posts from the control group, while from the diagnosed group, these numbers are 61,745 and 17,247, respectively. Vocab size depends on how many posts from a user are taken (1000, 3000, or 5000), but it is approximately 60,000 on average.

### *5.2 Role of Number of Users in Classification Accuracy*

This experiment has been carried out to determine the optimal number of cases required for an appropriate neural network. We have considered the different numbers of users to check how many users are at least required for the classification purpose. We have taken 150, 350, and 550 diagnosed users' posts. A middle range of the number of posts is taken for feature extraction for every user is 3000 (at most) posts. Also, the versions of the dataset manipulated in the experiment are balanced datasets, i.e., for 150 diagnosed users, 150 control users' posts are taken. Figure 1 presents the depression classification accuracy of different learning models on datasets with varying number of users.

### *5.3 Role of Number of Posts per User in Classification Accuracy*

The reason for considering the number of posts for a user employed is due to the prejudice that learning algorithms' give efficient results in most of the cases with somewhat bigger datasets. Some previous research suggests that people with depression have a low footprint in social activities or more activities at night. While creating the dataset, we observe that people with diagnosis statements have approximately 15% fewer posts than the control group. We have also found that a low number of posts (taken 1000 posts per user) or too many (taken 5000 or more) for a user-

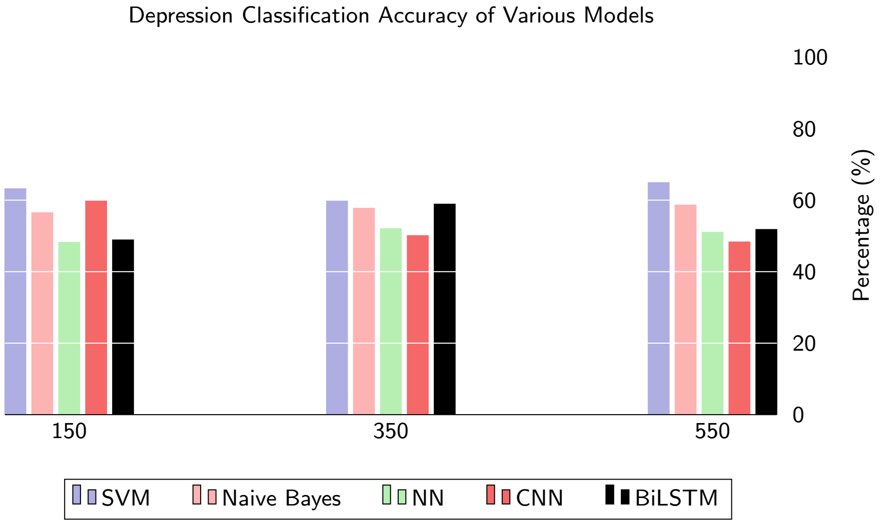


**Fig. 1** Role of number of users in classification accuracy

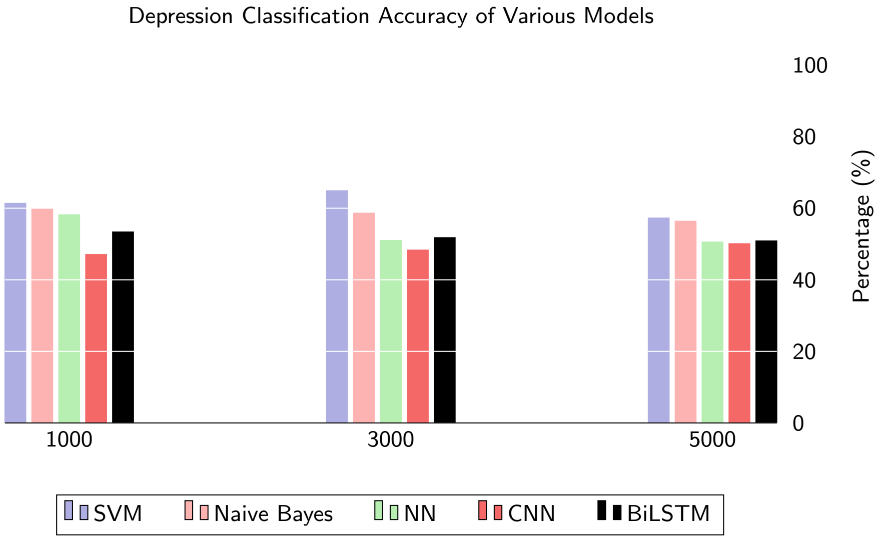


**Fig. 2** Role of number of posts per user in classification accuracy

produced somewhat low accuracy. We found little impact of number of posts per user on the classification accuracy observable Fig. 2. Three versions (1000 posts per user, 3000 posts per user, and 5000 posts per user) of balanced datasets are considered for this experiment.

## 6 Research Agenda

We propose to use bidirectional peephole BiLSTM at one level and attention layer at another level (Fig. 5). Before providing user posts as input to the hierarchical neural network, these are converted into tensors by passing them through the embedding layer. The embedding layer is initialized in two ways: by pre-trained (Word2Vec) word embedded vectors or other values. For the second option, word tensors are created using a vocabulary formed from a training corpus of posts. We compared the results of pre-trained Word2Vec and specially trained embedding for the purpose.

Compared to a feed forward neural network, recurrent neural network has a memory unit that remembers the last hidden state and considers it with the current input to produce output. However, the problem with simple RNN is the inability to remind long-standing memory. So long short-term memory (LSTM) is introduced to operate the situation. An LSTM unit (shown in Fig. 3) includes three layers to help its work: forget gate layer, input gate, and update gate. The forget gate layer removes information no longer required from the previous cell state, while the input gate layer decides what new information to add and puts it in the cell with forget layer output.

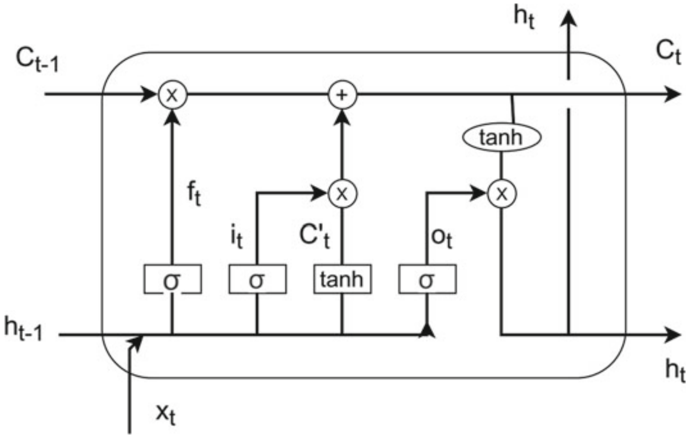


**Fig. 3** LSTM unit

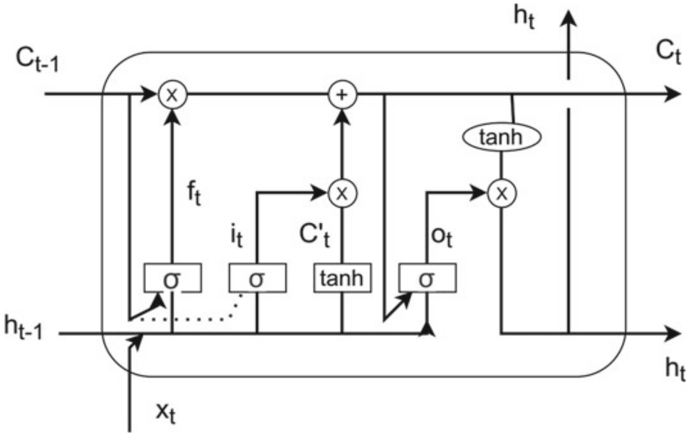


**Fig. 4** Peephole LSTM unit

Finally, the update gate layer decides what will remain on the cell state and provide as output. Bidirectional units are taken to examine previous inputs and future inputs as part. Peephole LSTM variant has little difference while deciding what to forget, input, and update. Every layer is provided with the cell state (as visible in Fig. 4) and the current input, which provides a performance advantage to the whole neural network. The following equations are for peephole LSTM, it can be noticed from the $C_t$ factor in $f_t$, $i_t$, and $o_t$

$$f_t = \sigma(W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f), \tag{1}$$

$$i_t = \sigma(W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i), \tag{2}$$

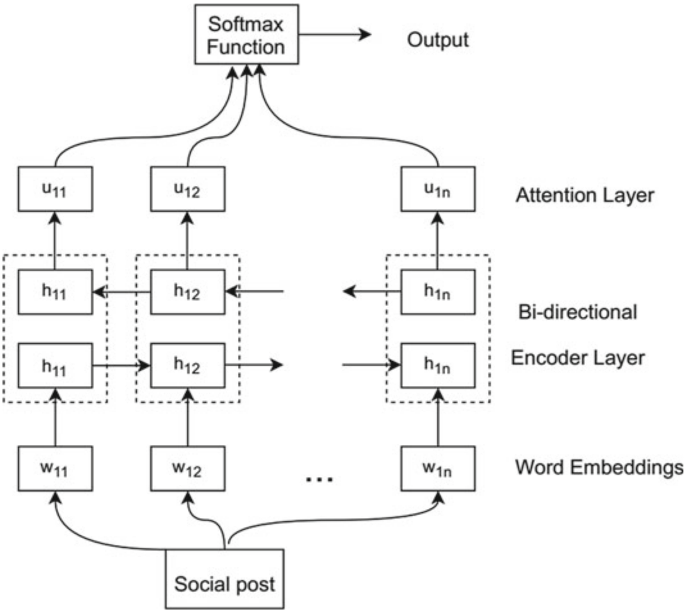


**Fig. 5** Bidirectional peephole LSTM with attention layer

$$o_t = \sigma(W_o \cdot [C_{t-1}, h_{t-1}, x_t] + b_o), \tag{3}$$

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{4}$$

$$h_t = o_t \cdot \tanh(C_t). \tag{5}$$

From Fig. 3, ft represents the output of the forget layer, this will be joined with previous cell state $C_t - 1$ for determining what values to be thrown away and what to keep for next cell state. At the input gate layer, it decides what new information will be added on the cell state while $C'_t$ creates candidate values. ot is calculated at updation layer to get the next hidden state from the current cell state. All the three gates make use of current input $x_t$, previous hidden state $h_t - 1$. Besides these two values forget and input gates take $C_t - 1$ as third input while update gate takes $C_t$. Examination of $C_t - 1$ and $C_t$ in decision-making by these gates is the peephole mechanism which is not in a simple LSTM unit (Fig. 3).

Output from the BiLSTM layer is used to input the attention layer placed above it. The attention layer provides different priority to the parts of a sentence by varying weight assignments. It gives extra context for decision-making and helps identify the crucial parts of posts focused by the neural network while producing output (Fig. 5).

# 7 Conclusion

From experiments, we learn that though the size of the dataset is an essential aspect in the domain, it does not affect our results that much. Also, different machine learning and deep learning algorithms have been applied on varying dataset versions constructed with the different number of control and diagnosed users. We tried to build a standard dataset resembling the one followed by various authors in previous works. We have followed the dataset creation procedure to build a standard dataset for one of the reputed workshops in the domain. Even then, we cannot ignore the fact that there may be some result variation if the mental health domain experts have prepared the dataset.

# References

1. The world health report 2001: Mental disorders affect one in four people. https://www.who.int/news/item/28-09-2001-the-world-health-report-2001-mental-disorders-affect-one-in-four-people. Accessed: 2022-01-12
2. Depression and Other Common Mental Disorders (2017) Global health estimates. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf. Accessed: 2022-01-30
3. Coronavirus data & statistics. https://www.cdc.gov/datastatistics/index.html. Accessed: 2022-01-12
4. Bhatia S, Kesarwani Y, Basantani A, Jain S (2021) Engaging smartphones and social data for curing depressive disorders: an overview and survey. In: Proceedings of the international conference on paradigms of computing, communication and data sciences. Springer, pp 809–820
5. Dalal S, Jain S (2021) Smart mental healthcare systems. In: Web semantics. Elsevier, pp 153–163
6. Dalal S, Jain S, Dave M (2019) A systematic review of smart mental healthcare. In: Proceedings of the 5th international conference on cyber security & privacy in communication networks (ICCS)
7. Basantani A, Kesarwani Y, Bhatia S, Jain S (2021) Emocure: utilising social media data and smartphones to predict and cure depression. IOP Conf Ser: Mater Sci Eng 1110:012010
8. Althouse BM, Allem JP, Childers MA, Dredze M, Ayers JW (2014) Population health concerns during the United States' great recession. Am Jo Prev Med 46(2):166–170
9. Ayers JW, Althouse BM, Allem JP, Rosenquist JN, Ford DE (2013) Seasonality in seeking mental health information on Google. Am J Prev Med 44(5):520–525
10. Hassanpour S, Tomita N, DeLise T, Crosier B, Marsch LA (2019) Identifying substance use risk based on deep neural networks and Instagram social media data. Neuropsychopharmacology 44(3):487–494
11. Xue Y, Li Q, Jin L, Feng L, Clifton DA, Clifford GD (2014) Detecting adolescent psychological pressures from micro-blog. In: International conference on health information science. Springer, pp 83–94
12. Gkotsis G, Oellrich A, Velupillai S, Liakata M, Hubbard TJ, Dobson RJ, Dutta R (2017) Characterisation of mental health conditions in social media using informed deep learning. Sci Rep 7(1):1–11
13. Lin H, Jia J, Guo Q, Xue Y, Huang J, Cai L, Feng L (2014) Psychological stress detection from cross-media microblog data using deep sparse neural network. In: 2014 IEEE International conference on multimedia and expo (ICME). IEEE, pp 1–6

14. Lin H, Jia J, Guo Q, Xue Y, Li Q, Huang J, Cai L, Feng L (2014) User-level psychological stress detection from social media using deep neural network. In: Proceedings of the 22nd ACM international conference on multimedia, pp 507–516
15. Trotzek M, Koitka S, Friedrich CM (2018) Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences. IEEE Trans Knowl Data Eng 32(3):588–601
16. Trotzek M, Koitka S, Friedrich CM (2018) Word embeddings and linguistic metadata at the clef 2018 tasks for early detection of depression and anorexia. In: CLEF (working notes)
17. Ahmad H, Asghar MZ, Alotaibi FM, Hameed IA (2020) Applying deep learning technique for depression classification in social media text. J Med Imaging Health Inf 10(10):2446–2451
18. Cong Q, Feng Z, Li F, Xiang Y, Rao G, Tao C (2018) Xa-bilstm: a deep learning approach for depression detection in imbalanced data. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, pp 1624–1627
19. Ive J, Gkotsis G, Dutta R, Stewart R, Velupillai S (2018) Hierarchical neural model with attention mechanisms for the classification of social media text related to mental health. In: Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic, pp 69–77
20. Sekulić I, Strube M (2020) Adapting deep learning methods for mental health prediction on social media. arXiv preprint arXiv:2003.07634
21. Uban AS, Rosso P (2020) Deep learning architectures and strategies for early detection of self-harm and depression level prediction. In: CEUR workshop proceedings, vol 2696. Sun SITE Central Europe, pp 1–12
22. Preoţiuc-Pietro D, Sap M, Schwartz HA, Ungar L (2015) Mental illness detection at the world well-being project for the clpsych 2015 shared task. In: Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, pp 40–45
23. Ren L, Lin H, Xu B, Zhang S, Yang L, Sun S et al (2021) Depression detection on reddit with an emotion-based attention network: algorithm development and validation. JMIR Med Inf 9(7):e28754
24. Sadeque F, Xu D, Bethard S (2017) Uarizona at the clef erisk 2017 pilot task: linear and recurrent models for early depression detection. In: CEUR workshop proceedings, vol 1866. NIH Public Access
25. Uddin AH, Bapery D, Arif ASM (2019) Depression analysis from social media data in bangla language using long short term memory (lstm) recurrent neural network technique. In: 2019 International conference on computer, communication, chemical, materials and electronic engineering (IC4ME2). IEEE, pp 1–4
26. Coppersmith G, Dredze M, Harman C (2014) Quantifying mental health signals in twitter. In: Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, pp 51–60