Lecture Notes in Networks and Systems 955

Umakanta Nanda
Asis Kumar Tripathy
Jyoti Prakash Sahoo
Mahasweta Sarkar
Kuan-Ching Li *Editors*

Advances in Distributed Computing and Machine Learning

Proceedings of ICADCML 2024, Volume 1

Springer

Lecture Notes in Networks and Systems

Volume 955

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

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

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

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

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

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

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

Umakanta Nanda · Asis Kumar Tripathy ·
Jyoti Prakash Sahoo · Mahasweta Sarkar ·
Kuan-Ching Li
Editors

Advances in Distributed Computing and Machine Learning

Proceedings of ICADCML 2024, Volume 1

Springer

Editors
Umakanta Nanda
School of Electronics Engineering
VIT-AP University
Amaravati, Andhra Pradesh, India

Jyoti Prakash Sahoo
Department of Computer Science
and Information Technology
Institute of Technical Education
and Research (ITER)
Siksha 'O' Anusandhan (Deemed to be
University)
Bhubaneswar, Odisha, India

Kuan-Ching Li
Department of Computer Science
and Information Engineering (CSIE)
Providence University
Taichung, Taiwan

Asis Kumar Tripathy
School of Information Technology
and Engineering (SITE)
Vellore Institute of Technology University
Vellore, Tamil Nadu, India

Mahasweta Sarkar
Department of Electrical and Computer
Engineering
San Diego State University (SDSU)
San Diego, CA, USA

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-97-1840-5 ISBN 978-981-97-1841-2 (eBook)
https://doi.org/10.1007/978-981-97-1841-2

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

Paper in this product is recyclable.

Preface

This volume of Lecture Notes in Networks and Systems is dedicated to the 5th International Conference on Advances in Distributed Computing and Machine Learning (ICADCML 2024). The ICADCML 2024 is an annual forum that aims to bring together ideas, innovations, lessons, etc. associated with distributed computing and machine learning, and their application in diverse areas. Distributed computing performs an increasingly important role in recent computing trends along with machine learning (ML). This conference aspires to stimulate research in Advances in Distributed Computing and Machine Learning. The conference was organized by the School of Electronics and Engineering, VIT-AP University, India, in Collaboration with College of Engineering, San Diego State University, USA, during January 05–06, 2024. Since 2020, ICADCML conference series has been a global platform for computer science researchers to exchange research results and ideas on the foundations and applications of distributed computing and machine learning.

With a history of 37 years of innovation in educational and research domain, VIT has been a forerunner in delivering quality education. Consistently ranked among the top educational institutes in the country, the VIT group of institutions have had a proud tradition of pursuing knowledge and excellence. In keeping with this tradition, the leadership at VIT-AP resonates with a dynamic blend of academic initiative and industry partnership with a vision of creating one of the finest academic destinations in the world. The VIT-AP campus, which is poised to become one of the country's best campuses, offers several avenues to explore your interests, identify core competencies, and engage in an evolving lifecycle of education and growth.

Since its inception in 1961, the College of Engineering, San Diego State University, has evolved into a top choice for serious engineering students and researchers from all over the globe. College of Engineering, SDSU, delivers a broad-spectrum, world-class engineering education, combined with practical research experience. The College of Engineering, SDSU, is dedicated to innovative education, discovery, and dissemination of knowledge. SDSU equips students to think with internal perspective, to design solutions that meet human and societal needs, and to create economic value that helps sustain the San Diego region and beyond.

We would like to convey our earnest appreciation to all the authors for their contributions to this book. We would like to extend our gratitude to all the reviewers for their constructive comments on all papers, especially we would like to thank the organizing committee for their hard work. Finally, we would like to thank the Springer publications for producing this volume.

Amaravati, India — Umakanta Nanda
Vellore, India — Asis Kumar Tripathy
Bhubaneswar, India — Jyoti Prakash Sahoo
San Diego, USA — Mahasweta Sarkar
Taichung, Taiwan — Kuan-Ching Li

Contents

Editors and Contributors

About the Editors

Umakanta Nanda is a distinguished academician in the field of Electronics Engineering. He received the MTech and Ph.D. degrees in Electronics and Communication Engineering from the National Institute of Technology, Rourkela, India, in 2010 and 2017, respectively. He is currently working as an Associate Professor and Dean of the School of Electronics Engineering at VIT-AP University, India. He has more than 15 years of teaching and research experience in different educational institutions. He has guided more than 20 UG and PG student projects. 5 Ph.D. scholars have received their Doctorate degree and another 4 scholars are working under him in areas like Analog and Mixed-signal integrated circuits, beyond CMOS devices and circuits, application-specific processor design, and embedded systems design. He has published more than 90 research papers, including reputed SCI and SCOPUS-indexed journals, conference proceedings, and book chapters. He is also co-inventor of 6 patents which have been published. He has successfully conducted many workshops, FDPs, STTPs, seminars, value-added courses, training programs, and conferences. He has also worked as a reviewer and editor of numerous journals and conferences.

Asis Kumar Tripathy is a professor in the School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India. He has more than ten years of teaching experience. He completed his Ph.D. from the National Institute of Technology, Rourkela, India, in 2016. His areas of research interests include wireless sensor networks, cloud computing, Internet of things and advanced network technologies. He has several publications in refereed journals, reputed conferences and book chapters to his credit. He has served as a program committee member in several conferences of repute. He has also been involved in many professional and editorial activities. He is a senior member of IEEE and a member of ACM.

Jyoti Prakash Sahoo is an experienced Assistant Professor and a Senior Member of IEEE, currently working at the Department of Computer Science and Information Technology, Institute of Technical Education and Research, Siksha ’O’ Anusandhan (Deemed to be University). He previously worked as an Assistant Professor with C. V. Raman College of Engineering, Bhubaneswar (now C. V. Raman Global University). He is an active member of several academic research groups, including the Scalable Adaptive Yet Efficient Distributed (SAYED) Systems Group at the Queen Mary University of London, the Intelligent Computing and Networking (ICN) Research Group at East China Normal University, and the Modern Networking Lab at National Taiwan University of Science and Technology. He has expertise in the field of Edge Computing and Machine learning. He also serves several journals and conferences as an editorial or reviewer board member. He served as Publicity Chair, Web Chair, Organizing Secretary, and Organizing Member of technical program committees for many national and international conferences. Being a WIPRO Certified Faculty, he has also contributed to industry-academia collaboration, student enablement, and pedagogical learning.

Mahasweta Sarkar is currently working as a professor of the Department of Electrical and Computer Engineering and senior associate dean, Global Campus at San Diego State University. Her M.S. and Ph.D. degrees were completed at the University of California, San Diego (UCSD), in 2003 and 2005, respectively. She received her B.S. degree in Computer Science and Engineering (Summa Cum Laude) in May 2000 from San Diego State University. Dr. Sarkar is a recipient of the “President’s Leadership Award for Faculty Excellence” for the year 2010. She delivered invited lectures and keynotes at different universities spread all over the globe. The talks were on wireless body area networks and brain–computer interfaces. Her research interest lies in the area of MAC layer power management algorithms and quality-of-service issues and protocols in WLANs, WMANs, WBANs, sensor networks and wireless ad-hoc networks.

Kuan-Ching Li is currently appointed as a distinguished professor at Providence University, Taiwan. He is a recipient of awards and funding support from several agencies and high-tech companies and also received distinguished chair professorships from universities in several countries. He has been actively involved in many major conferences and workshops in program/general/steering conference chairman positions and as a program committee member and has organized numerous conferences related to high-performance computing and computational science and engineering. Professor Li is the editor-in-chief of technical publications *Connection Science* (Taylor & Francis), *International Journal of Computational Science and Engineering* (Inderscience) and *International Journal of Embedded Systems* (Inderscience) and serves as an associate editor, editorial board member and guest editor for several leading journals. Besides publication of journal and conference papers, he is the co-author/co-editor of several technical professional books published by CRC Press, Springer, McGraw-Hill and IGI Global.

Contributors

T. S. Aarthi Radha Department of ECE, Saveetha Engineering College, Chennai, India

Alhassan Abdulhamid School of Computer Science, AI, and Electronics, University of Bradford, Bradford, UK

D. P. Acharjya School of Computer Science and Engineering, VIT, Vellore, India

K. Adinkrah-Appiah Sunyani Technical University, Brong-Ahafo, Ghana

Agarwal Shikha Department of Bioengineering and Biotechnology, Birla Institute of Technology, Mesra, India

Adil Ahamed Techno International New Town, Kolkata, West Bengal, India

V. Anitha National Engineering College, Kovilpatti, Tamil Nadu, India

M. R. Archana Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India

Lakshmi Bandikolla School of Electronics Engineering, VIT-AP University, Vijayawada, Andhra Pradesh, India

Neha Bangar E&TC Engineering, MIT Academy of Engineering, Pune, India

Francesco Benedetto P4TE- Signal Processing for Telecommunications and Economics Laboratory, Economics Department, University of ROMA TRE, Rome, Italy

Vaishali Bodade Fr. C. Rodrigues Institute of Technology, Navi Mumbai, India

R. M. Bommi Center for System Design, Chennai Institute of Technology, Chennai, India

M. Krishna Chaitanya School of Electronics Engineering, VIT-AP University, Amaravati, India

Krishnakant Chaubey Department of Electronics & Communication Engineering, National Institute of Technology Patna, Patna, India

Avigyan Chowdhury Techno International New Town, Kolkata, West Bengal, India

Asish Kumar Dalai VIT AP University, Amaravati, AP, India

Jerry Daniel CIG CDAC, Trivandrum, India

Amit Dash Odisha University of Technology, Bhubaneswar, India

Arabinda Dash Silicon Institute of Technology, Bhubaneswar, India

Lalit Deore E&TC Engineering, MIT Academy of Engineering, Pune, India

Vivek R. Dhanade Fr. C. Rodrigues Institute of Technology, Navi Mumbai, India

Manu Elappila CHRIST (Deemed to be University), Bengaluru, India

Muhammed Fuhad CHRIST (Deemed to be University), Bengaluru, India

P. Ganesan Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

Vinayak Garudi Fr. C. Rodrigues Institute of Technology, Navi Mumbai, India

Kaivalya Gauns National Institute of Technology Goa, Ponda, Goa, India

Demissie Jobir Gelmecha Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia

Ketema Adere Gemeda Department of CSE and Center for Electrical System and Electronics, Adama Science and Technology University, Adama, Ethiopia

Stenin K. George CHRIST (Deemed to be University), Bengaluru, India

Ibrahim Ghafir School of Computer Science, AI, and Electronics, University of Bradford, Bradford, UK

Chayan Ghosh Techno International New Town, Kolkata, West Bengal, India

Krishnendu Ghosh Techno International New Town, Kolkata, West Bengal, India

M. P. S. Guhan Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India

M. Harini National Engineering College, Kovilpatti, Tamil Nadu, India

P. Infant Vinoth Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India

Akanksha Jaiswal Electronics and Communication Engineering, NIT Hamirpur, Hamirpur, Himachal Pradesh, India

Jaiverdhan Department of ECE, National Institute of Technology, Srinagar, Uttarakhand, India

P. A. Jayasri PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India

Sachin Malayath Jose University of Liverpool, Liverpool, UK

Ch. Sai Jyothi Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

Sohag Kabir School of Computer Science, AI, and Electronics, University of Bradford, Bradford, UK

Aniket Kamble E&TC Engineering, MIT Academy of Engineering, Pune, India

S. Kanagamalliga Department of ECE, Saveetha Engineering College, Chennai, India

Ishan Rajesh Kasat School of Information Technology and Engineering, VIT, Tamil Nadu, Vellore, India

Manjur Kolhar Department of Computer Science, College of Arts and Science, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia

Addapalli V. N. Krishna CHRIST (Deemed to be University), Bengaluru, India

Aman Kumar Electronics and Communication Engineering, NIT Hamirpur, Hamirpur, Himachal Pradesh, India

Jayendra Kumar School of Electronics Engineering, VIT-AP University, Inavolu, Amaravathi, Andhra Pradesh, India

Nancy Kumari School of Computer Science and Engineering, VIT Bhopal University, Bhopal, India

C. S. Lajitha CIG CDAC, Trivandrum, India

Aniruddh Lawande National Institute of Technology Goa, Ponda, Goa, India

Ci Lei School of Computer Science, AI, and Electronics, University of Bradford, Bradford, UK

L. M. I. Leo Joseph Department of Electronics and Communication Engineering, SR University, Warangal, Telangana, India

Yan Ma College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China

Santosh Kumar Maharana Odisha University of Technology and Research, Bhubaneswar, India

Sambit Mallick Department of Electronics and Communication Engineering, Heritage Institute of Technology, Kolkata, West Bengal, India

Songa Manikanta Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

V. Manimaran National Engineering College, Kovilpatti, Tamil Nadu, India

R. Manimegalai PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India

A. Mary Joy Kinol SIMATS School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

Federica Massimi P4TE- Signal Processing for Telecommunications and Economics Laboratory, Economics Department, University of ROMA TRE, Rome, Italy

Rani Rachel Mathew Amity Private School, Sharjah, United Arab Emirates

Abdul Kayom Md Khairuzzaman School of Electronics Engineering, VIT-AP University, Vijayawada, Andhra Pradesh, India

Vikas Mehetre Fr. C. Rodrigues Institute of Technology, Navi Mumbai, India

Subhadarshini Mohanty Odisha University of Technology and Research, Bhubaneswar, India

Hitesh Mohapatra School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

Subasish Mohapatra Odisha University of Technology and Research, Bhubaneswar, India

Apurv Mule Fr. C. Rodrigues Institute of Technology, Navi Mumbai, India

D. A. Murali Krishnan School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

P. Muthu Krishnammal School of Electronics Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Hariharan Muthusamy Department of ECE, National Institute of Technology, Srinagar, Uttarakhand, India

Shamanth Nagaraju CHRIST (Deemed to be University), Bengaluru, India

A. Nagaswathy Rathnavel (RVS), College of Arts and Science, Sulur, Tamil Nadu, India

D. Nagendra Kumar Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India

Vinita Nair Department of Computer Science, Gujarat University, Ahmedabad, India

P. Narmatha National Engineering College, Kovilpatti, Tamil Nadu, India

Wesene Derbe Negeri Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia

Chandan Panda Odisha University of Technology, Bhubaneswar, India

Trilochan Panigrahi National Institute of Technology Goa, Ponda, Goa, India

Anurag Parajuli Department of CSE, CMR Institute of Technology, Bengaluru, India

Jyoti Pareek Department of Computer Science, Gujarat University, Ahmedabad, India

H. Parveen Sultana School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

Snigdha Paul Department of Electronics and Communication Engineering, Heritage Institute of Technology, Kolkata, West Bengal, India

Somya Prakash Department of Computer Science and Information Technology, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India; Department of Electrical and Electronics Engineering, Siksha O Anusandhan Deemed To Be University, Bhubaneswar, Odisha, India

N. Prasanna Kumar School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

Jagdeep Rahul Department of Electronics & Communication Engineering, Rajiv Gandhi University, PaPum Pare, India

Harsh Raja Department of Computer Science and Information Technology, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India; Department of Computer Science and Engineering, Siksha O Anusandhan Deemed To Be University, Bhubaneswar, Odisha, India

S. Rajalingam Department of EEE, Saveetha Engineering College, Chennai, India

Sriramachandran Ramesh School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

Smita Rath Department of Computer Science and Information Technology, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

Davinder Singh Rathee Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia

Kanta Rathee All India Jat Heroes' Memorial College (AIJHM), Rohtak, India

Vandana Reddy CHRIST (Deemed to be University), Bengaluru, India

C. S. Reshmah PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India

Sanchari Saha Department of CSE, CMR Institute of Technology, Bengaluru, India

Seemanti Saha Department of Electronics & Communication Engineering, National Institute of Technology Patna, Patna, India

Kshira Sagar Sahoo Department of Computing Science, Umeå University, Umeå, Sweden

K. L. Sailaja Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

G. Sajiv Saveetha School of Engineering, SIMATS, Chennai, India

S. Santhana Hari Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India

Dibyasha Sarangi Odisha University of Technology, Bhubaneswar, India

B. S. Sathish Department of Computer Science Engineering AI, Sri Venkateshwara College of Engineering, Bengaluru, India

S. Selvavarshini National Engineering College, Kovilpatti, Tamil Nadu, India

S. Kalai Selvi National Engineering College, Kovilpatti, Tamil Nadu, India

Anindya Sen Department of Electronics and Communication Engineering, Heritage Institute of Technology, Kolkata, West Bengal, India

Rupesh Kumar Shah Department of CSE, CMR Institute of Technology, Bengaluru, India

Divya Sharma Department of Electronics and Communication Engineering, Ajay Kumar Garg Engineering College, Lucknow, India

Lakhan Dev Sharma School of Electronics Engineering, VIT-AP University, Inavolu, Amaravathi, Andhra Pradesh, India

M. M. Sharma Department of ECE, Malviya National Institute of Technology, Jaipur, India

K. Sharmila Banu Vellore Institute of Technology, Vellore, India

Priyanka Singh School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India

Ram Sewak Singh Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia

Rakesh Kumar Sinha Department of Bioengineering and Biotechnology, Birla Institute of Technology, Mesra, India

V. G. Sivakumar Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

Dhruv Umesh Sompura School of Information Technology and Engineering, VIT, Tamil Nadu, Vellore, India

R. Sridhar Department of ECE, Saveetha Engineering College, Chennai, India

M. Suganya Rathnavel (RVS), College of Arts and Science, Sulur, Tamil Nadu, India

B. Tejaswari Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India

Lijo Thomas CIG CDAC, Trivandrum, India

S. Thulasi Prasad Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

Anurag Tripathy Kijiji Canada, Toronto, ON, Canada

B. K. Tripathy School of Information Technology and Engineering, VIT, Tamil Nadu, Vellore, India

G. Uganya Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Hans Upadhyay Department of Computer Science and Information Technology, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, India

S. Vaishnavi PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India

S. Vengadakrishnan Department of ECE, Saveetha Engineering College, Chennai, India

Priya Verma Department of ECE, Malviya National Institute of Technology, Jaipur, India

Usha Verma E&TC Engineering, MIT Academy of Engineering, Pune, India

Comparative Analysis of Deep Learning-Based Hybrid Algorithms for Liver Disease Prediction

Dhruv Umesh Sompura, B. K. Tripathy, Anurag Tripathy, and Ishan Rajesh Kasat

Abstract The liver is the largest and one of the most important organs inside the human body and its disorder affects the functioning of vital activities of the body. However, India alone has more than a million people diagnosed with liver diseases each year; hence, it is important to detect them at an early stage. We need an automatized process that can predict the symptoms of liver diseases and this can be done using machine learning algorithms or more specifically Deep Neural Network (DNN) algorithms. While these algorithms are used by researchers, their accuracies are low in the context of the security of humans being involved in the process. In this paper, we try to propose hybrid deep learning algorithms, which have higher degrees of accuracy. We have used three hybrid algorithms: CNN combined with LSTM (99.02%), CNN combined with GRU (98.38%), and CNN combined with RNN (99.48%). Experimental results establish that these accuracies, especially in the model where CNN is combined with RNN, are much higher than those of the existing algorithms in the literature. The methods used to achieve the results which involve preprocessing the data, feature selection, training, and testing followed by the use of the hybrid models, and then finally, the results obtained are analyzed using various metrics.

Keywords Liver disease · CNN · LSTM · GRU · RNN

D. U. Sompura · B. K. Tripathy (✉) · I. R. Kasat
School of Information Technology and Engineering, VIT, Tamil Nadu, Vellore 632014, India
e-mail: tripathybk@vit.ac.in

D. U. Sompura
e-mail: dhruvumesh.sompura2020@vitstudent.ac.in

I. R. Kasat
e-mail: ishanrajesh.kasat2020@vitstudent.ac.in

A. Tripathy
Kijiji Canada, 1, York Street, Suite 1010, Toronto, ON M5J 0B6, Canada

U. Nanda et al. (eds.), *Advances in Distributed Computing and Machine Learning*,
Lecture Notes in Networks and Systems 955,
https://doi.org/10.1007/978-981-97-1841-2_1

1 Introduction

The liver, the largest internal organ of the body, carries out essential biological processes such as creating glycogen, bile secretion, triglyceride and cholesterol production, and blood clotting factors. Usually, more than 75% of the liver's tissue is compromised before a decline in liver function becomes serious. Therefore, it is essential to start treating illnesses as soon as they are discovered in order to prevent them from getting worse.

Deep learning (DL) is a subtopic of machine learning that mimics the human brain to process data by integrating data to create accurate predictions [1]. Incorporating the concept into the artificial neural network context, the Deep Neural Networks' (DNNs) models [1] have been framed. The human brain follows the deep learning process in its cognitive processes. There are several special types of DNNs; Convolutional Neural Networks (CNNs) [2], Recurrent Neural Networks (RNNs) [3], Generative Adversarial Neural Networks [4], and so on. DNN models have been found to be efficient in image processing and classification [5, 6]. Also, there are several other applications of DNN models. DNN models have more hidden layers than normal ANNs [7]. The number of hidden layers is to be selected judiciously so that at the same time it does not increase the complexity and improves accuracy. Deep learning algorithms rely on large amounts of data and work multiple times, always tweaking to improve results. Deep learning techniques have become very important in health care nowadays for disease prediction from medical datasets [8–12]. Classification of data is an important activity in data mining. Among various DL techniques, classification algorithms are widely used in disease prediction. There are several popular machine learning algorithms like: K-Nearest Neighbors (KNNs), Support Vector Machine (SVM), Random Forest (RF), Naive Bayes Classifiers (NBCs), etc. Hybrid models are usually more efficient than individual components if the strong points of the component models are taken into account while making the hybridization. We have studied the same and made some points on it. It is our aim in this work to form hybrid DNN-based algorithms and find their predictive accuracy in order to suggest the most suitable one for practical applications.

2 Literature Review

The study conducted by [13] employed machine learning techniques such as Logistic Regression (LR), K-Nearest Neighbors (KNNs), Support Vector Machines (SVMs), Decision Trees (DTs), Naive Bayes (NB), and Random Forest (RF). The performance of these algorithms was evaluated and compared using several metrics, including F1-score, recall, precision, and accuracy. The objective of the study was to minimize the expenses associated with liver disease diagnosis [14] by employing machine learning algorithms to accurately classify patients as either liver patients or non-liver patients.

Logistic Regression is employed in order to do a comparative analysis of the findings vis-à-vis those of other scholars. The results of the study indicated that Logistic Regression exhibited superior accuracy compared to the Naïve Bayes Classifier, Decision Tree, Support Vector Machine (SVM), Artificial Neural Network (ANN), and K-Nearest Neighbors' (KNNs) algorithms. The study used the Naïve Bayes and Support Vector Machine (SVM) classifier algorithms to predict and analyze liver disease, as described in Ref. [15]. The text highlights the presence of two primary factors that play a significant role in comprehending the appropriateness of the separate methodologies: the duration required to carry out the prediction process and the precision of the predictive outcomes. The superiority of the SVM classification algorithm is evident due to its consistently high accuracy rates. However, in terms of the execution time of the predictive process, the Naive Bayes classifier demonstrates more applicability as it has the shortest execution time. In Ref. [16], Decision Tree methods were employed for the purpose of classification and thereafter evaluated in terms of seven performance indicators, including accuracy percentage (ACC%), mean absolute error (MAE), precision (PRE), recall (REC), F-measure (FME), Kappa Statistics, and runtime. The algorithms employed in the study included J48, LMT, Random Tree, Random Forest, REPTree, Decision Stump, and Hoeffding Tree. The analysis demonstrates that the Decision Stump technique yields superior accuracy compared to other techniques. The study conducted in Ref. [17] employs a software engineering methodology that incorporates categorization and feature selection techniques. This study examines the performance of six classification methods, namely J48, Random Forest, Logistic Regression, SMO (Support Vector Machine), IBk (k closest Neighbor), and Naïve Bayes, in relation to the Indian Liver Patient Dataset (ILPD). The process of creating intelligent liver disease prediction software (ILDPS) involves the utilization of feature selection and classification prediction algorithms, which are based on a software engineering model. The present study centers on the creation of software designed to facilitate the prediction of disease severity by analyzing a range of symptoms.

The methodology suggested in the aforementioned study [18] involves constructing classification models for the purpose of forecasting the fate of liver disease. The construction of the model involves the utilization of four distinct phases. The initial step is the use of the first min–max normalization method on the original liver patient datasets obtained from the UCI repository. In the second phase, the utilization of Particle Swarm Optimization (PSO) for feature selection results in the acquisition of a subset of the liver patient dataset. This subset is produced from the entire normalized liver patient datasets and consists solely of attributes that are deemed significant. In the third phase, the dataset is subjected to categorization algorithms. During the fourth phase, the accuracy will be determined by calculating the root mean square value and the root mean error value. The J48 method is widely regarded as a high-performing algorithm, particularly when used in conjunction with Particle Swarm Optimization (PSO) for feature selection. Ultimately, the assessment is conducted utilizing accuracy metrics. The results of the experimental research indicate that the J48 algorithm exhibits superior performance compared to all other categorization algorithms. The application of a genetic algorithm [19] was employed to optimize the boosted C5.0 algorithm, resulting in the generation of rules specifically tailored for the detection and diagnosis of liver illness. The established principles for liver disease diagnosis are deemed ideal and accurate when compared with other offered methodologies. The application of a genetic algorithm has resulted in a significant reduction in the number of rules, decreasing from an initial count of 92 to a final count of 24. The statistical metrics indicate that the suggested method exhibits superior performance compared to boosted C5.0, and it also requires a considerable amount of time for diagnosis. In a previous study, a range of classification algorithms including Naïve Bayes, Decision Tree, Multi-layer Perceptron, and k-NN were examined and evaluated [20]. A comparison examination reveals that Naïve Bayes demonstrates superior precision compared to other algorithms. However, in terms of criteria such as recall and sensitivity, Logistic Regression and Random Forest have advantages over other algorithms. The research conducted by the authors in Ref. [21] examined many algorithms, including C4.5, Naive Bayes, Decision Tree, Support Vector Machine, Backpropagation Neural Network, and Classification and Regression Tree Algorithms. It has been observed that the C4.5 algorithm yields superior outcomes in comparison to alternative algorithms. In the future, it is possible to develop an enhanced version of the C4.5 algorithm by incorporating different parameters. The authors of the study introduced a liver disease prediction system named Modified Convolutional Neural Network-based liver disease prediction system (MCNN-LDPS) in Ref. [16]. The purpose of this system was to achieve precise liver disease prediction outcomes. Upon investigation, it is evident that the MCNN-LDPS approach demonstrates superior outcomes in terms of heightened accuracy and precision. The present study involved a comparative comparison of the performance of the research method under investigation and the established multilayer perceptron neural network (MLPNN). The issue of spatial invariance of the input data sample in relation to the lack of ability has been addressed in this study by the integration of a genetic algorithm with a Convolutional Neural Network (CNN).

3 Proposed Algorithm

3.1 Existing Methods

- Classified liver disease using artificial neural networks' (ANNs) classification algorithm resulted in low accuracy.
- The testing accuracies of the MLP, SMO and Logistic Regression method were found to be 77.54%, 71.36% and 74.36% respectively.
- We are going to use the following hybrid deep learning algorithms for our work and compare their accuracies.

3.2 LSTM-CNN

Numerous methods exist for boosting model performance, including alterations to batch size, epochs, dataset refinement, adjustments to the training, validation, and test dataset ratios, modification of loss functions, and varying model architectures, among others. In this study, we aim to enhance model performance by specifically modifying the model's architecture. Our focus is on assessing whether the CNN-LSTM model can offer superior predictions of liver disease cases when compared to the LSTM model. The CNN layers are responsible for feature extraction from the input data, while the LSTM layers facilitate sequence prediction.

3.3 CNN + GRU

To leverage the strengths of the Gated Recurrent Unit (GRU) module, adept at handling time sequence data, and the CNN module, a hybrid neural network called GRU-CNN was introduced. The inputs comprise time series data from the energy system and information from the spatiotemporal matrix, with the goal of predicting future load values. The CNN component excels at processing two-dimensional data, including spatiotemporal matrices and images. It employs local connections and shared weights to directly extract local features from spatiotemporal matrix data, efficiently representing them through convolution and pooling layers. The CNN module consists of two convolution layers, each followed by pooling, and the resulting high-dimensional data is flattened to one-dimensional data before being fed into a fully connected layer. Conversely, the GRU module's purpose lies in capturing long-term dependencies. It accomplishes this by learning valuable information from historical data through memory cells over extended periods, allowing it to forget irrelevant information over time. The GRU module takes time series data as input and comprises multiple gate recursion units, with their outputs connected to fully connected layers. Ultimately, the load prediction result is obtained by averaging across all neurons within the fully connected layer.

3.4 CNN + RNN

The proposed model harnesses the strengths of both Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM). Initially, a Conv1D CNN layer is utilized to process input vectors, focusing on extracting local features found at the text level. The output of this CNN layer, which comprises feature maps, is subsequently fed into the RNN layer, consisting of LSTM units or cells. Here, the RNN layer capitalizes on the local features obtained from the CNN layer, enabling the model to grasp and model long-term dependencies within these local features.

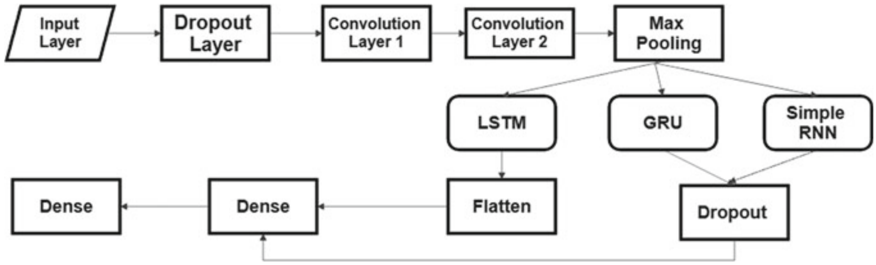

Fig. 1 Various layers for hybrid deep learning algorithms

By combining CNN's proficiency in capturing local patterns and LSTM's ability to capture sequential dependencies, the proposed model achieves a balanced approach to understanding and utilizing information in the input data. This approach is particularly valuable in tasks that require a comprehensive understanding of both short-term nuances and long-term relationships within the data. The interplay between CNN and LSTM layers enhances the model's capability to analyze complex sequences effectively, making it a powerful tool in various applications, such as natural language processing and time series analysis.

3.5 Diagrammatic Representation of Algorithms

Figure 1 shows the flowchart of the hybrid deep learning algorithms. The common steps in the start include the input layer (to bring initial data in the input) followed by the dropout layer (used to reduce overfitting in the model) followed by the two convolution layers (as these are CNN hybrid algorithms). This is followed by max pooling (creates a down-sampled feature map) in all cases. Now for LSTM, it is followed by a flattened layer (converts the feature map to a single vector), while GRU and simple RNN are followed by another dropout layer. The last two layers are both dense layers for all three algorithms.

4 Methodology

4.1 Architecture

The architecture of the code is explained further ahead (Fig. 2).

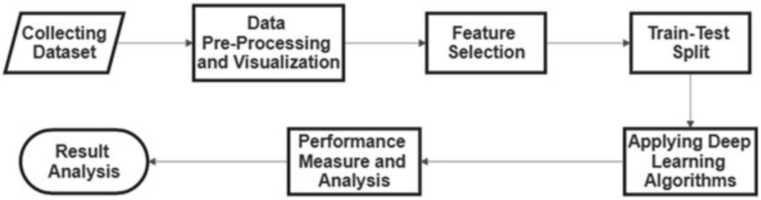

Fig. 2 Architecture diagram

4.2 Dataset Description and Sample Data

Data Set Information

The important details regarding the dataset are provided below (UCI Machine Learning Repository, 2013). The dataset was collected from the UCI Machine Learning Repository. The total number of records in this dataset for patients with liver conditions is 416, while the total number of records for patients without liver conditions is 167. The "Dataset" column acts as a class label to divide people into two groups: those with liver illness and those without. The data comes from North East of Andhra Pradesh, India. There are 142 female patients and 441 male patients included in the patient records.

Attribute Information

- Age of the patient.
- The gender of the patient.
- Total Bilirubin.
- Direct Bilirubin.
- Alkaline Phosphatase.
- Alanine Aminotransferase.
- Aspartate Aminotransferase.
- Total Proteins.
- Albumin.
- Albumin and Globulin ratio.
- Class: field used to bifurcate the data into two subsets, distinguishing patients with liver disease from those without.

Before loading the dataset, we should import all the required libraries such as pandas, tokenizer, numpy, seaborn, and label encoder to perform operations of implementing deep learning models as well to perform steps of data preprocessing. Here, we have downloaded the dataset from the UCI repository and saved it as indian_liver_patient.csv which is now loaded and can be read as a data frame which is now named as data.

4.3 Data Preprocessing and Visualization

While creating our paper, the dataset that we imported from the repository was not clean and formatted, and before employing the deep learning models on the data, it is very necessary to clean and put formatted data; hence, data preprocessing is required, and it is basically the process of preparing the raw data and making it ready for the deep learning model.

Observations

Using the "data.describe" command, we may get a number of insights from the dataset, including:

- Ten characteristics and one output variable make up the dataset.
- Notably, the "Albumin and Globulin ratio" feature has four missing data.
- While all other properties are represented as numeric variables, gender is a non-numeric variable.
- The values of Alkaline Phosphatase, Alanine Aminotransferase and Aspartate Aminotransferase should be converted to floating-point values for improving accuracy (Table 1).

Filling of Missing Values

Finding missing variables and replacing them with the mean values are known as imputation. The "Albumin and Globulin ratio" column in our dataset had four missing values, which were replaced with the column's mean value of 94.7. Subsequently, it is clear from the second image that the "A/G ratio" column is no longer blank.

Table 1 Dataset values used for the calculation along with their data types

Variable	Data type
Age	int64
Total_Bilirubin	float64
Direct_Bilirubin	float64
Alkaline_Phosphotase	int64
Alamine_Aminotransferase	int64
Aspartate_Aminotransferase	int64
Total_Protiens	float64
Albumin	float64
Albumin_and_Globulin_Ratio	float64
Dataset	int64
dtype	object

Identifying Duplicate Values

There were about 13 duplicates found. It is crucial to keep in mind that duplicate values in a medical dataset can still be accurate; thus, we decided against eliminating any of these duplicate items.

Resampling

It is necessary to be applied due to the imbalancing nature of the dataset which is reflected in more number of liver disease patients than the non-liver patients in the dataset. For this an oversampling technique called SMOTE is applied, which generates new values for the minority data and synthesizes new samples. This process improves accuracy when implementing machine learning algorithms using the Weka Tool. Additionally, PCA was applied to achieve better results. Finally, combination of SMOTE and PCA was used to compare accuracy among various machine learning algorithms.

4.4 Feature Selection

Feature selection is a process of figuring out which inputs are the best for the model and checking if there is a possibility of eliminating certain inputs. Considering the dataset, we can see a very high linear relationship between Total and Direct Bilirubins, and by considering this linear relationship, Direct Bilirubin can be opted to be dropped, but as per medical analysis, Direct Bilirubin constitutes almost 10% of the Total Bilirubin and this 10% may prove crucial in obtaining higher accuracy for the model; thus, none of the features are removed.

4.5 Train–Test Split

A DL algorithm is evaluated using this. The process entails splitting the dataset into two subsets. It is a quick and simple process to carry out. We have taken into account 80% of the training data and 20% of the testing data for the liver disease prediction model.

4.6 Code Implementation

The code for the three hybrid algorithms, RNN + CNN, LSTM + CNN, and GRU + CNN, was run on an IDE and the results were noted for further analysis. The outputs for the hybrid algorithms listed the total parameters, trainable parameters and non-trainable parameters along with the loss and accuracy at each epoch.

5 Experimentation

5.1 Result Analysis

In the existing method [10], the researchers have used MLCNN-LDPS which provided an accuracy of 90.75%. We have used three hybrid algorithms: CNN + LSTM (99.02%), CNN + GRU (98.38%), and CNN + RNN (99.48%) and have achieved accuracy as high as 99.48% using filters like upscaling and PCA. We also used various algorithms and got the following accuracies—Naïve Bayes: 76%, Random Forest: 80.26%, Logistic: 72%, SVM: 76.93%, KNN: 76.67%. We used PCA and SMOTE to increase the number of cases in the dataset in a balanced way. This gave us better accuracies (Fig. 3; Table 2).

Figure 4 shows the heat maps for the dataset. This helps visualize which areas matter the most in the dataset. The data variables are the x and y axes of the map.

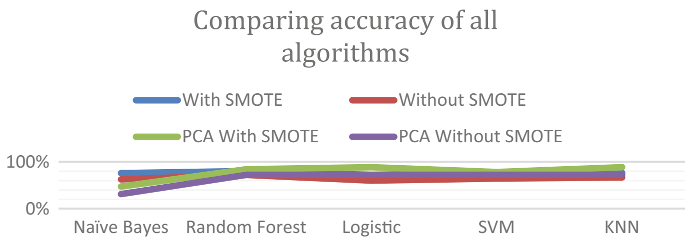

Fig. 3 Graphical representation of the table above

Table 2 Table contains the accuracy of all the algorithms run on weka before and after applying PCA and SMOTE

Algorithms	With SMOTE (%)	Without SMOTE (%)	PCA with SMOTE (%)	PCA without SMOTE (%)
Naïve Bayes	76	62.6072	47.4667	31.38
Random forest	80.2667	72.0412	84.2667	72.3842
Logistic regression	72	60.2058	88.5333	72.3842
SVM	76.9333	64.494	78.5333	72.3842
KNN	76.6667	66.7238	88.5333	72.3842