DEEP LEARNING:
THEORY, ARCHITECTURES AND APPLICATIONS IN SPEECH, IMAGE AND LANGUAGE PROCESSING

Editors:
Gyanendra Verma
Rajesh Doriya

Bentham Books

Deep Learning: Theory, Architectures and Applications in Speech, Image and Language Processing

Edited by

Gyanendra Verma
National Institute of Technology Raipur,
Raipur, India

&

Rajesh Doriya
National Institute of Technology Raipur,
Raipur, India

Deep Learning: Theory, Architectures and Applications in Speech, Image and Language Processing

Editors: Gyanendra Verma and Rajesh Doriya

ISBN (Online): 978-981-5079-21-0

ISBN (Print): 978-981-5079-22-7

ISBN (Paperback): 978-981-5079-23-4

© 2023, Bentham Books imprint.

Published by Bentham Science Publishers Pte. Ltd. Singapore. All Rights Reserved.

First published in 2023.

BENTHAM SCIENCE PUBLISHERS LTD.

End User License Agreement (for non-institutional, personal use)

This is an agreement between you and Bentham Science Publishers Ltd. Please read this License Agreement carefully before using the ebook/echapter/ejournal (**"Work"**). Your use of the Work constitutes your agreement to the terms and conditions set forth in this License Agreement. If you do not agree to these terms and conditions then you should not use the Work.

Bentham Science Publishers agrees to grant you a non-exclusive, non-transferable limited license to use the Work subject to and in accordance with the following terms and conditions. This License Agreement is for non-library, personal use only. For a library / institutional / multi user license in respect of the Work, please contact: permission@benthamscience.net.

Usage Rules:

1. All rights reserved: The Work is the subject of copyright and Bentham Science Publishers either owns the Work (and the copyright in it) or is licensed to distribute the Work. You shall not copy, reproduce, modify, remove, delete, augment, add to, publish, transmit, sell, resell, create derivative works from, or in any way exploit the Work or make the Work available for others to do any of the same, in any form or by any means, in whole or in part, in each case without the prior written permission of Bentham Science Publishers, unless stated otherwise in this License Agreement.
2. You may download a copy of the Work on one occasion to one personal computer (including tablet, laptop, desktop, or other such devices). You may make one back-up copy of the Work to avoid losing it.
3. The unauthorised use or distribution of copyrighted or other proprietary content is illegal and could subject you to liability for substantial money damages. You will be liable for any damage resulting from your misuse of the Work or any violation of this License Agreement, including any infringement by you of copyrights or proprietary rights.

Disclaimer:

Bentham Science Publishers does not guarantee that the information in the Work is error-free, or warrant that it will meet your requirements or that access to the Work will be uninterrupted or error-free. The Work is provided "as is" without warranty of any kind, either express or implied or statutory, including, without limitation, implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the results and performance of the Work is assumed by you. No responsibility is assumed by Bentham Science Publishers, its staff, editors and/or authors for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products instruction, advertisements or ideas contained in the Work.

Limitation of Liability:

In no event will Bentham Science Publishers, its staff, editors and/or authors, be liable for any damages, including, without limitation, special, incidental and/or consequential damages and/or damages for lost data and/or profits arising out of (whether directly or indirectly) the use or inability to use the Work. The entire liability of Bentham Science Publishers shall be limited to the amount actually paid by you for the Work.

General:

1. Any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims) will be governed by and construed in accordance with the laws of Singapore. Each party agrees that the courts of the state of Singapore shall have exclusive jurisdiction to settle any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims).
2. Your rights under this License Agreement will automatically terminate without notice and without the

need for a court order if at any point you breach any terms of this License Agreement. In no event will any delay or failure by Bentham Science Publishers in enforcing your compliance with this License Agreement constitute a waiver of any of its rights.

3. You acknowledge that you have read this License Agreement, and agree to be bound by its terms and conditions. To the extent that any other terms and conditions presented on any website of Bentham Science Publishers conflict with, or are inconsistent with, the terms and conditions set out in this License Agreement, you acknowledge that the terms and conditions set out in this License Agreement shall prevail.

Bentham Science Publishers Pte. Ltd.
80 Robinson Road #02-00
Singapore 068898
Singapore
Email: subscriptions@benthamscience.net

CONTENTS

FOREWORD

Machine Learning proved its usefulness in many applications in Image Processing and Computer Vision, Medical Imaging, Satellite imaging, Remote Sensing, Surveillance, *etc*., over the past decade. At the same time, Machine Learning, particularly Artificial Neural Networks has evolved and demonstrated excellent performance over traditional machine learning algorithms. These methods are known as Deep Learning.

Nowadays, Deep Learning has become the researcher's first choice in contrast to traditional machine learning due to its apex performance on speech, image, and text processing. Deep learning algorithms provide efficient solutions to problems ranging from image and speech processing to text processing. The research on deep learning is getting enriched day by day as we witness new learning models.

Deep learning models significantly impacted speech, image, and text-domain and raised the performance bar substantially in many standard evaluations. Moreover, new challenges are easily tackled by utilizing deep learning, which older systems could not have handled. However, it is challenging to comprehend, let alone guide, the learning process in deep neural networks; there is an air of uncertainty about exactly what and how these networks learn.

This book aims to provide the audience with a basic understanding of deep learning and its different architectures. Background knowledge of machine learning helps explore various aspects of deep learning. By the end of the book, I hope that the reader understands different deep learning approaches, models, pre-trained models, and gains familiarity with implementing various deep learning algorithms using multiple frameworks and libraries.

Dr. Shitala Prasad
Scientist, Institute for Infocomm Research, A*Star
Singapore 138632
Singapore

PREFACE

Machine Learning proved its usefulness in many applications in the domain of Image Processing and Computer Vision, Medical Imaging, Satellite imaging, Remote Sensing, Surveillance, *etc* ., over the past decade. At the same time, Machine Learning methods themselves have evolved, particularly deep learning methods that have demonstrated significant performance over traditional machine learning algorithms.

Today's Deep Learning has become researchers' first choice in contrast to traditional machine learning due to its apex performance in many applications in the domain of speech, image, and text processing. Deep learning algorithms provide efficient solutions to problems ranging from vision and speech to text processing. The research on deep learning is getting enriched day by day as we witness new learning models.

This book contains two major parts. Part one includes the fundamentals of Deep Learning, theory, and architecture of Deep Learning. Moreover, this part provides a detailed description of the theory, frameworks, and non-conventional approaches to deep learning. It covers foundational mathematics that is essential in understanding the framework. Moreover, it covers various kinds of models found in practice.

Chapter 1 contains the basic operating understanding, history, evolution, and challenges associated with deep learning. We will also cover some basic concepts of mathematics and the hardware requirements for deep learning implementation, and some of its popular software frameworks. We will start with neural networks, which focus on the basics of neural networks, including input/output layers, hidden layers, and how networks learn through forward and backpropagation. We will also cover the standard multilayer perceptron networks and their building blocks. Moreover, we will include a review of deep learning concepts in general and deep learning in particular to build a basic understanding of this book. Chapters 2–7 are based on applying artificial intelligence to medical images with various deep learning approaches. It also covers the application of Deep Learning in lung cancer detection, medical imaging, and COVID-19 analysis.

The second part, chapters 8–10, is dedicated to sentiment analysis using deep learning and machine learning techniques. This book section covers the experimentation and application of deep learning techniques and architectures in real-world applications. It details the salient approaches, issues, and challenges in building ethically aligned machines. An approach inspired by traditional Eastern thought and wisdom is also presented.

The third part, Chapters 11–15, is miscellaneous and covers the different artificial intelligence approaches used to explain the machine learning models that enhance transparency between the user and the model. A review and detailed description of the use of knowledge graphs in generating explanations for black-box recommender systems and elaborative education ecosystems for sustainable quality education is provided. Reinforcement learning is a semi-supervised learning technique for portfolio management.

Gyanendra Verma
National Institute of Technology Raipur
Raipur, India

&

Rajesh Doriya
National Institute of Technology Raipur
Raipur, India

List of Contributors

Aqib Ali Sayed	Amity School of Engineering and Technology, Amity University Mumbai, Maharashtra 410206, India
Abdul Jabbar Perumbalath	School of Computer Science, Mahathma Gandhi University,Kottayam, Kerala, India
Anil Verma	Department of Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India
Aman Singh	Department of Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India
Archana Mathiazhagan	School of Computing, SASTRA Deemed University, Thanjavur 613401, India
Babita Panda	School of Electrical Engg., KIIT University, Bhubaneswar, India
Chellapilla Vasantha Lakshmi	Dayalbagh Educational Institute, Agra, India
Chellapilla Patvardhan	Dayalbagh Educational Institute, Agra, India
Divya Anand	Department of Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India
Elakkiya Rajasekar	School of Computing, SASTRA Deemed University, Thanjavur 613401, India
Gyanendra Verma	Department of Information Technology, National Institute of Technology, Raipur, India
Harshawardhan Tiwari	Jyothy Institute of Technology, Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka, India
Harshee Pitroda	Department of Computer Engineering, NMIMS University, Mukesh Patel School of Technology Management & Engineering, Mumbai, India
Ishani Saha	Department of Computer Engineering, NMIMS University, Mukesh Patel School of Technology Management & Engineering, Mumbai, India
Jay Prajapati	Department of Data Science, SVKM's NMIMS, Mumbai, Maharashtra, India
Jayalakshmi Ramachandran Nair	Research Scholar, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India
Jaykumar Suraj Lachure	Department of Information Technology, National Institute of Technology, Raipur, India
Mayank Gupta	Department of Computer Science and Engineering, Punjab Engineering College, Chandigarh, India
Muhamed Ilyas Poovankavil	PG and Research Department of Computer Science, Sullamussalam Science College, Areekode, Malappuram Dt, Kerala, India
Manisha Tiwari	Department of Computer Engineering, NMIMS University, Mukesh Patel School of Technology Management & Engineering, Mumbai, India

Natarajan Balasubramanian	School of Computing, SASTRA Deemed to be University, Thanjavur, Tamilnadu 613401, India
Nehha Seetharaman	Amity School of Engineering and Technology, Amity University Mumbai, Maharashtra 410206, India
Poonam Saini	Department of Computer Science and Engineering, Punjab Engineering College, Chandigarh, India
Parth Kalkotwar	Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
Prajwal Sethu Madhav	Jyothy Institute of Technology, Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka, India
Priyanka Prashanth Kumar	Jyothy Institute of Technology, Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka, India
Rakesh Kumar Dhaka	ITM University, Gwalior, India
Rajkumar Narayanan	Department of Sciences, St. Claret College, Bengaluru, Karnataka, India
Rajalakshmi Elangovan	School of Computing, SASTRA Deemed University, Thanjavur 613401, India
Rishika Vij	Department of Veterinary Physiology & Biochemistry, Dr. GC Negi College of Veterinary & Animal Science, Palampur, Himachal Pradesh, India
Rajesh Doriya	Department of Information Technology, National Institute of Technology, Raipur, India
Saleena Thorayanpilackal Sulaiman	PG and Research Department of Computer Science, Sullamussalam Science College, Areekode, Malappuram Dt, Kerala, India
Sampurna Panda	ITMUniversity,Gwalior, sampurnapanda, India
Sanay Shah	Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
Saurav Tiwari	Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
Sindhu Nair	Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
Siba Panda	Department of Data Science, SVKM's NMIMS, Mumbai, Maharashtra, India
Sumathy Pichai Pillai	Department of Computer Science & Applications, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India
Srishti Sakshi Sinha	Department of CSE, Pondicherry University, Puducherry, India
Sukrati Chaturvedi	Dayalbagh Educational Institute, Agra, India
Sushila Ratre	Amity School of Engineering and Technology, Amity University, Mumbai, Maharashtra 410206, India
Trupthi Muralidharr	Jyothy Institute of Technology, Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka 560082, India
Vatsal Khandor	Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
Uma Vijayasundaram	Department of Computer Science, Pondicherry University, Puducherry, India

CHAPTER 1

Deep Learning: History and Evolution

Jaykumar Suraj Lachure[1,*], **Gyanendra Verma**[1] and **Rajesh Doriya**[1]

[1] *National Institute of Technology Raipur, Raipur, India*

Abstract: Recently, deep learning (DL) computing has become more popular in the machine learning (ML) community. In the field of ML, the most widely used computational approach is DL. It can solve many complex problems, cognitive tasks, and matching problems without any human performance or interface. ML cannot handle large amounts of data and DL can easily handle it. In the last few years, the field of DL has witnessed success in a range of applications. DL outperformed in many application domains, *e.g.*, robotics, bioinformatics, agriculture, cybersecurity, natural language processing (NLP), medical information processing, *etc.* Despite various reviews on the state of the art in DL, they all concentrated on a single aspect of it, resulting in a general lack of understanding. There is a need to provide a better beginning point for comprehending DL. This paper aims to provide a more comprehensive overview of DL, including current advancements. This paper discusses the importance of DL and introduces DL approaches and networks. It then explains convolutional neural networks (CNNs), the most widely used DL network type and subsequent evolved model starting with LeNET, AlexNet with the Letnet-5, AlexNet, GoogleNet, and ResNet networks, and ending with the High-Resolution network. This paper also discusses the difficulties and solutions to help researchers recognize research gaps for DL applications.

Keywords: Convolution neural network, Deep learning applications, Deep Learning, Image classification, Machine Learning, Medical image analysis.Natural Language Processing.

INTRODUCTION

In the last decade, machine learning (ML) models [1 - 3] have been widely used in every field and have been applied in versatile applications like classification, image/video retrieval, text mining, multimedia, anomaly detection, attack detection, video recommendation, image classification, *etc.* Nowadays, deep learning (DL) is frequently employed in comparison to other machine learning methods. DL stands for representative learning. The unpredictable expansion of

* **Corresponding author Jaykumar Suraj Lachure:** National Institute of Technology Raipur, India; E-mail: jaykuamrlachure@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

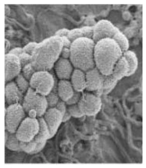
DL and distributed learning necessitates ongoing study. Deep and distributed learning studies are continuing to emerge as a result of unanticipated advances in data availability and huge advancements in hardware technologies such as High-Performance Computing (HPC). DL is a Neural Network (NN) that outperforms its predecessors. DL also employs transformations and graph technology to create multi-layer learning models. In fields such as Natural Language Processing (NLP), data processing, visual data processing, and audio and speech processing, the most recent DL techniques have achieved extraordinary performance. The representation of input data is often what determines the success of an ML approach. A proper data representation outperforms a poor data representation. Thus, for many years, feature engineering has been a prominent study topic in ML. This method helps to build features from raw data. It also involves a lot of human effort and is quite field-specific. These are the scale-invariant feature transform (SIFT), histogram of oriented gradients (HOG), and bag of words (BoW).

The DL algorithms automatically extract features, and this helps researchers extract discriminative features with minimal human effort and field knowledge. A multi-layer data representation architecture extracts low-level features at the first layer, while the last layer extracts high-level features. Artificial Intelligence (AI) is the basis of all technology, including ML, DL, and NLP, *etc.*, which processes data for particular applications, much like in the human brain's basic sensory regions. The human brain can automatically derive data representation using different scenes. This procedure's output is the classified objects, while the input is the incoming scene information. This mimics the human brain's workings. Thus, it accentuates DL's key advantage.

Due to its significant success, DL is presently one of the most important research fashions in ML. Architectures, issues, computational tools, the evolution matrix, and applications are all significant elements in DL. In DL networks, convolutional neural networks (CNN) are widely employed. CNN automatically finds key features, making it the most widely used. Therefore, we delved deep into CNN by showing its core elements. From the AlexNet network to the GoogleNet with high-resolution network, each uses the most prevalent CNN topologies.

Several deep learning models have solely dealt with one application or issue in recent years, such as examining CNN architectures or deep learning. There are different applications like autonomous machines, deep learning for plant disease detection and classification, deep learning for security and malicious attack detection, and so on. Table **1** shown below provides a few domains and applications of DL. Prior to diving into DL applications, it is important to grasp the concepts, problems, and benefits of DL. Learning DL to address research gaps and applications takes a lot of time and research. Our proposal is to conduct an extensive review of DL to provide a better starting point for a comprehensive grasp of DL.

Table 1. Different Domains of DL and Applications.

 **Internet &Cloud**	 **Medicine & Biology**	 **Media & Entertainment**	 **Security & Defense**	 **Autonomous Machines**	 **Agriculture**
Image Classification	Cancer Cell Detection	Video Captioning	Face Detection	Pedestrian Detection	Crop Recommendation
Speech Recognition	Diabetic Grading	Video Search	Video Surveillance	Lane Tracking	Leaf Disease Detection
Language Translation	Drug Discovery	Real Time Translation	Satellite Imagery	Recognize Traffic Sign	Fruit Classification
Language Processing	Drug-Drug Interaction	Recommendation	Malicious Attack	Object Detection	Smart Irrigation
Sentiment Analysis	Drug-protein Interaction	Image/ video Retrieval	Firewall Security	Object Tracking	Leaf Identification

For our review, we focused on open challenges, computational tools, and applications. This review can also be a springboard for further DL discussions.

The review helps individuals learn more about recent breakthroughs in DL research, which will help them grow in the field. In order to deliver precise alternatives to the field, researchers would be given greater autonomy. Here are our contributions:

- This review aids researchers and students in gaining comprehensive knowledge about DL.
- We will describe the historical overview of neural networks.
- We discuss deep learning approaches using Deep Feedforward Neural Networks, Deep Backward Neural Networks, and CNN, as well as their concepts, theories, and current architectures.
- We describe the different CNN architectures like AlexNet, GoogleNet, and ResNet.
- We describe deep learning models that use auto-encoders, long short-term memory, and a deep belief network architecture.

CHAPTER 2

Application of Artificial Intelligence in Medical Imaging

Sampurna Panda[1], **Rakesh Kumar Dhaka**[1] and **Babita Panda**[2,*]

[1] *ITM University, Gwalior, India*

[2] *School of Electrical Engineering, KIIT University, Bhubaneswar, India*

Abstract: The emergence of the Internet of Things (IoT) and Artificial Intelligence (AI) applications in many industries is due to recent developments in technology and connectivity. This paper outlines various industry initiatives in healthcare that utilize machine learning techniques. To meet this rising demand, considerable investment is required to develop new medical imaging algorithms, such as those that can be used to diagnose disease diagnostic systems errors, which can yield ambiguous medical treatments. Early disease in imaging is usually predicted by machine learning and deep learning algorithms. Imaging tools use machine learning and deep learning techniques to analyze early disease. Medical imaging is on the cutting edge of deep learning techniques, specifically the application of convolution neural networks. The supervised or unsupervised algorithms are applied to a dataset containing specific instances, and then the predictions are displayed. Machines and deep learning approaches are excellent for data classification and automated decision-making.

Keywords: Artificial Intelligence, Deep Learning, Internet of Things (IoT), Machine Learning, Neural Network.

INTRODUCTION

The computer system's machine learning algorithm is essential to improve its ability to make accurate predictions and make decisions. The area of study that teaches computers how to study without requiring them to be explicitly programmed is known as machine learning [1]. Deep learning is a group of machine learning, which enables systems to gain an understanding of the world in terms of a pinpoint of ideas [2]. The growing performance of neural networks propelled the growth of deep learning in computer vision.

* **Corresponding author Babita Panda:** School of Electrical Engineering, KIIT University, Bhubaneswar, India; E-mail: pandababita18@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

MACHINE-LEARNING

Machine learning describes methods that enable computers to learn on their own and solve problems. A mathematical model should be trained so that it can be fed with useful input data and then produce valuable results. Fig. (**1**) shows the Pictorial visualization of Machine Learning.

Fig. (1). Pictorial visualization of Machine Learning.

Machine learning models are provided with training data and optimized to yield accurate predictions. To put it simply, the models are developed with the goal of delivering predictions for new data with an unseen level of certainty. To estimate the generalization ability of a model, you use a separate dataset called the validation set and apply the model's predictions to that dataset as a form of feedback for tuning the model. The model is evaluated using test data against which it has been trained and fine-tuned to see how it will perform with unseen data. It is possible to roughly categorize machine learning methods according to how the models' input data is utilized during training. An agent is a built-in reinforcement learning through trial and error, while one is optimizing a particular objective function. The computer is assigned the task of learning on its own without guidance.

Fig. (**2**) shows the relationship between artificial intelligence (AI), Machine learning (ML), Artificial Neural Network(ANN), and Convolutional Neural Network(CNN). In a nutshell, grouping is the quintessential use case. Most of the today's machine learning is working with real-time algorithms, such as supervised learning. Here, a set of already labeled or annotated images is given to the computer. Then, it is challenged to assign correct labels to new, previously unseen datasets that are constructed according to the learned rules. This consists of a collection of input-output examples.

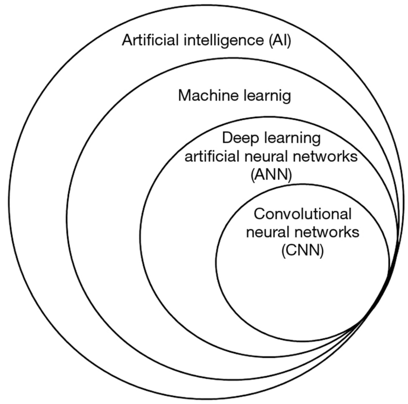


Fig. (2). Relation of AI with ML, ANN, CNN.

There are various machine learning techniques. Fig. (**3**) shows the different types of machine learning techniques available. In this section, different machine learning techniques are described, each with its process, advantages, and disadvantages.

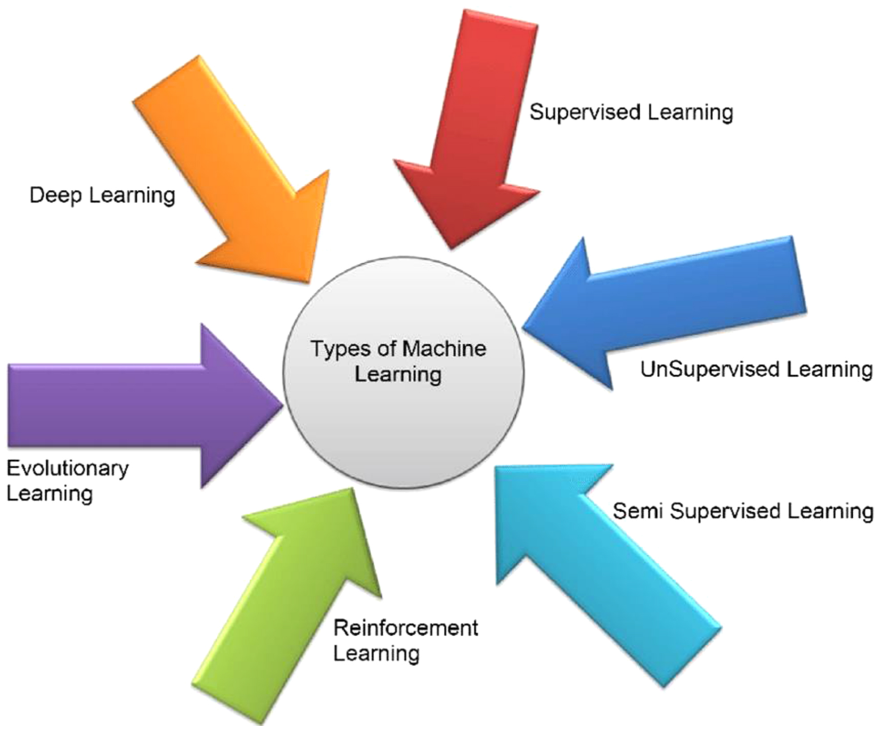


Fig. (3). Different types of machine learning techniques.

CHAPTER 3

Classification Tool to Predict the Presence of Colon Cancer Using Histopathology Images

Saleena Thorayanpilackal Sulaiman[1,*], Muhamed Ilyas Poovankavil[2] and **Abdul Jabbar Perumbalath[3]**

[1] *Sullamussalam Science College, Areekode, Malappuram, Kerala, India*

[2] *PG and Research Department of Computer Science, Sullamussalam Science College, Areekode, Malappuram Dt, Kerala – India*

[3] *School of Computer Science, Mahathma Gandhi University, Kottayam, Kerala, India*

Abstract: The proposed model compares the efficiency of CNN and ResNet50 in the field of digital pathology images. Deep learning methods are widely used in all fields of disease detection, diagnosis, segmentation, and classification. CNN is the widely used image classification algorithm. But it may show less accuracy in case of complex structures like pathology images. Residual Networks are a good choice for pathology image classification because the morphology of digital pathology images is very difficult to distinguish. Colon cancer is one of the common cancers, and it is one of the fatal diseases. If early-stage detection has been done using biopsy results, it will decrease the mortality rate. ResNet50 is selected among the variants as its computational complexity is moderate and provides high accuracy in classification as compared to others. The accuracy metric used here is the training and validation accuracy and loss. The training and validation accuracy of ResNet50 is 89.1% and 90.62%, respectively, whereas the training loss and validation loss are 26.7% and 24.33%, respectively. At the same time, for CNN, the accuracy is 84.82% and 78.12% and the loss is 36.51% and 47.33% .

Keywords: Colon cancer, CNN, H&E stained histopathology, ResNet50.

INTRODUCTION

There are nineteen different types of cancer that can affect a healthy person [1]. Among them, colon cancer is the third leading cancer in the US. Tumors inside of a patient's body are most commonly detected through medical imaging techniques, such as radiology. Different modalities like X-rays [2], CT scan, PET

* **Corresponding author Saleena Thorayanpilackal Sulaiman:** Sullamussalam Science College, Areekode, Malappuram, Kerala, India; E-mail: tssaleena@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

scan or MRI, *etc.,* can be used to find the position and rate of progress of cancer and thereby leading to classifying whether it is benign or malignant. But the confirmation can be done only through biopsy. So pathologists play a crucial role in disease diagnosis and treatment plan decision in case of cancer.

The workload of pathologists is very hectic as their number is very low as compared to their work.

Colon cancer originates from the large intestine, where the colon is the end of the digestive tract [3, 4]. The signs and symptoms showing for this cancer are very non-specific. So screening like colonoscopy in people of any age can reduce the mortality rate as this is a preventable and curable disease in the early stage. Pathological tests will be very beneficial in such situations.

Most colon adenocarcinomas are moderately differentiated, but some are poorly differentiated. So the classification may not be correct looking at histopathology images. So now, ancillary studies, or mainly ImmunoHistochemical staining is the prevalent method for the same [5 - 7]. It is a protein-based test that is very costly as compared with histopathology image analysis.

How colon cancer looks like and where it originated are shown in Fig. (**1**).

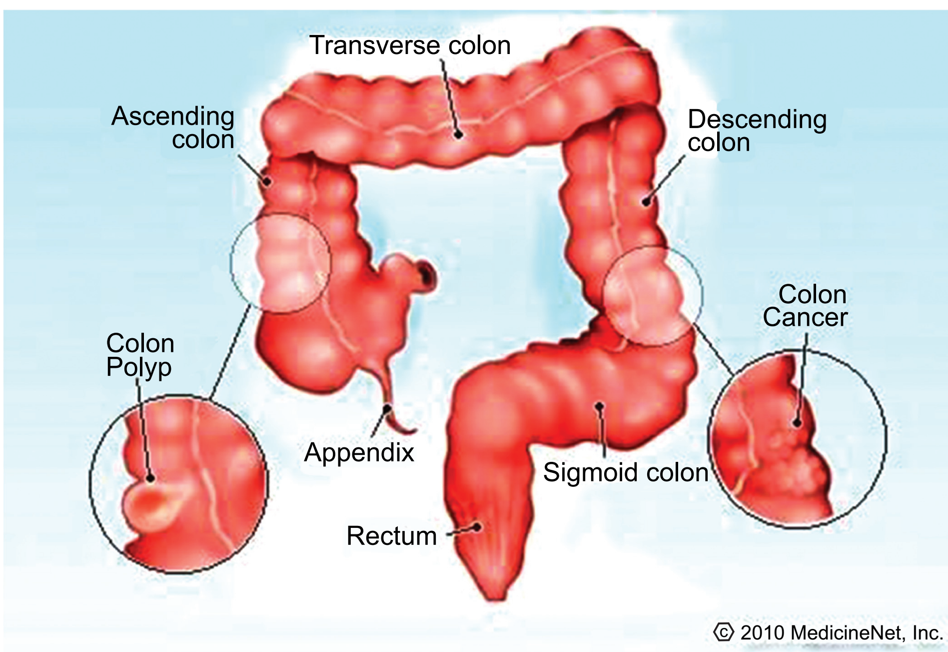


Fig. (1). Colon cancer-affected region [1].

Deep learning algorithms are very effective in [8] stain normalization [9], segmentation, classification, grading, staging, prognosis, *etc.,* especially neural networks. Digital pathology can be used as a good sort of alternative to the conventional glass slide-keeping process. It is easy to transport and can store the image without physical damage. Convolutional Neural Network (CNN) is a powerful and accurate alternative to digital image analysis and image classification methods. Computer vision can identify and classify the region of cancer present in the histopathology images which are very difficult using the naked eye. First, we have to train the model with adequate data and the best network should be chosen on a trial-and-error basis. [10] Pacal *et al.* described various CNN algorithms applied in colon cancer for various applications. Recently published 135 papers have been reviewed in that paper. Recently published 135 papers have been reviewed in that paper. Digital images of H&E stained histopathology images are used as input for these algorithms. The size and shape of the nucleus, the shape of colorectal polyps, accumulation of irregular tissues, *etc.,* are some of the features to be considered for the detection and classification of colon cancer.

Even though CNN is a powerful tool in image classification, the classification of pathology images is more difficult due to its morphology. Different powerful pre-trained networks are available for image classification. Such networks are working with the principle of transfer learning, and most of them are trained with ImageNet; and their feature extraction ability and their weights can be applied to our local dataset to create our own models. Only the last dense layer will be changed, and all other layers will be kept untouched.

This paper comprises a comparison of the CNN algorithm with ResNet50 [11] using a dataset of colon cancer obtained from an online repository. ResNet50 is one of the efficient and widely used variants of Residual Networks. It is a pre-trained network that has been trained upon the ImageNet dataset. This study proves that the ResNet50 is more appropriate for such complex image classification. The CNN that is used here has four convolutional layers followed by a pooling layer. A similar kind of CAD has been developed for the classification of breast cancer [12], Gastric and Colonic Epithelial Tumors [13], and ten other kinds of cancers [14, 15]. Nucleus detection models are also developed for several cancers [16, 17]. Noise removal is the major part of the pre-processing technique that can improve the efficiency of the model. As we are using the processed data in this model, no such methods are used here.

As direct data collection is the most tedious and almost hard-to-reach step in AI, especially in the case of medical data, public repositories are a blessing to some extent. Among several repositories, TCGA is one of them. The Cancer Genome

CHAPTER 4

Deep Learning For Lung Cancer Detection

Sushila Ratre[1,*], **Nehha Seetharaman**[1] and **Aqib Ali Sayed**[1]

[1] *Amity School of Engineering and Technology, Amity University Mumbai, Maharashtra 410206, India*

Abstract: By detecting lung cancer in advance, doctors can make the right decision to treat patients to ensure that they live long and healthy lives. This research aims to build a CNN model using a pre-trained model and functional API that would classify if a person had lung cancer or not based on a CT scan. This research uses CT scan images as input for the prediction model from the LUNA16 [Luna Nodule Analysis 2016] dataset for experimenting by using ResNet 50 and VGG 16. ResNet50 showed slightly high accuracy on test data compared to VGG16, which is 98%.

Keywords: ResNet 50, VGG 16, CNN, Lung Cancer, Deep Learning.

INTRODUCTION

Lung cancer is considered the world's most lethal cancer, taking countless lives each year. This is the most common cancer in the world, as well as the most common cause of death. This cannot be ignored and leads to death if not treated on time. It is a state that causes cancerous cells to split uncontrollably in the lungs. This causes the growth of tumors that impair a person's ability to breathe. According to World Health Organization (WHO) in 2019, lung cancer was the second leading cause of death earlier than 70 years in a hundred and twelve countries and ranked third or fourth in a further 23 countries among 183 countries across the globe [1]. Propitiously, early detection of cancer can significantly enhance survival rates. Lung tumor identification is done using numerous imaging techniques such as Computed Tomography Scans (CT), Sputum Cytology, Chest Lung X-rays, and Magnetic Resonance Imaging (MRI). Discernment means classifying a tumor into two cancerous tumors or non-cancerous tumors. An artificial intelligence approach is being used in this research to identify anomalies in lung CT scans using tumor-specific characteristics.

* **Corresponding author Sushila Ratre:** Amity School of Engineering and Technology, Amity University Mumbai, Maharashtra 410206, India; E-mail: suratre@mum.amity.edu

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

There are four main stages in non-small cell lung cancers [NSCLC], from stage 1 to stage 4. If cancerous cells are detected in Stage 1 in the lungs, and it has not di spersed outside to any other body part, then a person's survival rate is close to approximately five years. The tumor size should be less than 3 cm for Stage 1. Similarly, in stage 2, the size of the tumor is between 3-5 cm, and the survival rate is between 2-3 years [2]. The details regarding the stages and survival rate are given below in Table **1**.

Table 1. Stages of Lung Cancer [NSCLC].

S. No	Stage No	Size (In cm)	Dispersion	Rate of Survival (Approx.)
1	Stage 1	< 3 cm	Inside lung	5 Years
2	Stage 2	3-5 cm	Inside lung and nearby lymph nodes	3 Years
3	Stage 3	5-7 cm	Middle of the chest	1-2 Years
4	Stage 4	>7 cm	Both lungs and distant organs	0-6 Months

From (Fig. **1**), we can identify the location and position of the tumor that CT scan images can easily capture. Researchers can prepare a dataset according to the stages of lung cancer to train their model in the future.

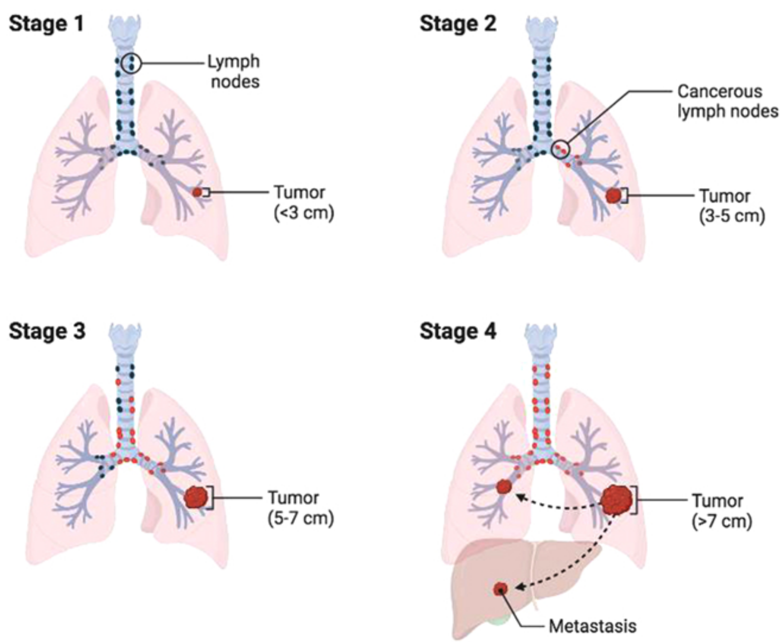


Fig. (1). Types of Stages of Non-Small Cell Lung Cancers [3].

According to the Global survey from the Source [4]: GLOBOCAN 2020, the new lung cancer cases is 2,206,771, and the death rate is very high, *i.e.,* 1,796,144. It is the motivation behind these research activities. Research has considered a total of 36 types of cancers, and lung cancer is the second deadliest disease for humans.

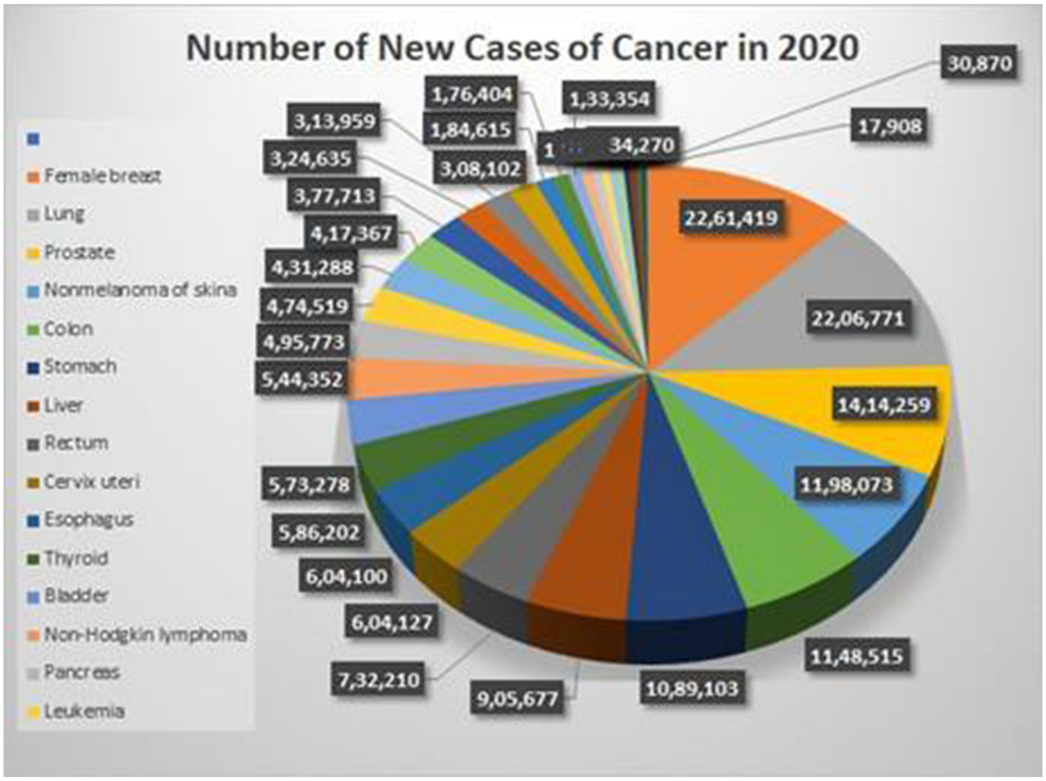


Fig. (2). Number of new cases of cancer in 2020 [4].

RELATED WORKS

QingZeng Song, Lei Zhao, XingKe Luo, and XueChen Dou, 2017 [1] used Lung Image Database Consortium (LIDC) - Image Database Resource Initiative (IDRI) dataset and compared three different categories of deep network models, namely CNN, DNN & SAE that are designed for lung calcification and passed the data to all the three networks, and leaky ReLU activation function is used. They concluded that the CNN network achieved the best performance with an accuracy of 84.15%, specificity of 84.15%, and sensitivity of 84.32%, which is the best among the three networks.

Siddharth Bhatia *et al.,* 2018 [5] have used an LIDC-IDRI dataset and deep residual learning. Features are extracted from CT scans using UNet and Resnet

CHAPTER 5

Exploration of Medical Image Super-Resolution in terms of Features and Adaptive Optimization

Jayalakshmi Ramachandran Nair[1,*], **Sumathy Pichai Pillai**[2] and **Rajkumar Narayanan**[3]

[1] *Research Scholar, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India*

[2] *Department of Computer Science & Applications, Bharathidasan University, Tiruchirappalli, Tamil Nadu India*

[3] *Department of Sciences, St. Claret College, Bengaluru, Karnataka India*

Abstract: Medical image processing takes many steps to capture, process, and convert the images for further analysis. The images are susceptible to distortions due to various factors related to the analysis tools, environment, system-generated faults, and so on. Image enhancement deals with enhancing the quality and resolution of images for accurately analyzing the original information from the images. The primary motivating aspect of research and reconstruction of such high-quality images and their challenges is image super-resolution for image upgrading. This chapter focuses on various image-enhancing strategies in implementing the super-resolution process. In this work, the methodologies of various image-enhancing strategies are explained clearly to provide the parameter selection points, feature comparisons, and performance evaluations that apply to high-resolution image processing. The drawbacks and challenges of each strategy are discussed to investigate the effectiveness of the methodologies. Further research is explored to find hybrid methods on various deep learning architectures to achieve higher accuracy in the field of medical image super-resolution.

Keywords: Image enhancement, Image feature mapping, Image processing, Image registration, Image resolution, Image segmentation, Medical imaging.

INTRODUCTION

Image enhancement improves the quality of pixel contents of the raw captured image data before sending it for further processing. Commonly used image enhancement techniques include contrast adjustment, spatial filtering, *etc.* The uniformity clears the uneven portion of pixel alignments and improves pixel qual-

* **Corresponding author Jayalakshmi Ramachandran Nair:** Research Scholar, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India; E-mail: jayabinoy2020@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

ity. Spatial noise removal filters improve the naturally occurring lines, dead zones in the pixels, and shear zones.

On the other hand, density intervals are aligned by evenly spreading the grey tone to achieve high enhancement without breaking the pixels. Image contrast can be enhanced using linear transformation techniques that expand the original grey level [1].

In many medical applications, there has been a requirement to renovate low-resolution images into high-resolution ones to identify the faults correctly so that the issue can be recognized and evaluated excellently. In image super-resolution, the lower-resolution images (LR) are transformed into high-resolution images (HR) [2, 3]. The main applications of image super-resolution are in surveillance and healthcare. In the case of surveillance, detecting and identifying the face on low-quality camera images is one of the needs [4]. The adjustment and enhancement of pixels enable the process of finding the matching pattern [5, 12]. The adaptation of lower resolution to higher resolution can be processed by handling the pixels one by one or a patch of pixels together [6, 13]. When captured using MRI systems, medical images are high quality in nature [2, 3, 7, 8]. In this case, the time to complete the full scan, the spatial coverage, and signal-to-noise ratio (SNR) factors become tricky [9]. The super-resolution principle helps generate high-quality enhanced images from low-quality images [1, 10, 11]. This will highly reduce the processing time.

LITERATURE REVIEW

The interaction of clinical imaging, extraction, transformation, and capacity requires manual mediation [12, 13]. The images acquired by this interaction are, for the most part, degraded and crumbled because of the impact of different noticeable factors, for example, vulnerabilities present in the extreme climate, types of equipment utilized in the clinical imaging measure, and so on [14, 15]. Decay in the resolution of clinical images alludes to the deficiency of pertinent and noticeable data in clinical images. To upgrade the nature of these images, an image goal improvement method, for example, image super-resolution (SR) innovation, is utilized [16, 17]. This procedure will recreate the LR and change them into HR, which is broadly used in applications like clinical imaging, image and face acknowledgment, observation, satellite imaging, *etc* [18, 19]. Late headways in clinical image resolutions have recommended the execution of profound learning methods to work on the nature of images. Deep learning strategies have accomplished critical significance in image-preparing applications. Profound learning calculations have a hearty portrayal capacity contrasted with ordinary procedures [20, 21]. SR imaging strategies utilizing convolutional neural network (CNN) procedures have given agreeable outcomes in working on the goals of clinical images [22].

In recent years, innovative adaptations of CNN approaches like Fast Super-Resolution Convolutional Neural Networks (FSRCNN) and Super-Resolution Convolutional Neural Networks (SRCNN) have been proposed for clinical imaging [23, 24]. There have been various attempts to extricate resolution issues of a clinical image with the help of FSRCNN and SRCNN by quite a few researchers.

METHODOLOGIES

The low-quality pixels of the images are enhanced to high-resolution images based on the degradation function denoted by D. If the high-resolution image and low-resolution image are considered, the peak amount of noise present in the low-resolution image is represented by σ.

$$I_{High} = D\,(I_{Low};\, \sigma) \tag{1}$$

The degradation parameter and σ are variables that deviate depending on the image quality, captured environment lighting, *etc.* Algorithms such as a neural network reveal the transfer function by inverting the degradation ratio using high-resolution and low-resolution data. Various super-resolution techniques are pre-up sampling super-resolution, post-up sampling super-resolution, residual networks, multi-stage residual networks, recursive networks, progressive reconstruction, multi-branch models, attention-based networks, and generative models. The efficiency of the super-resolution technique in producing the best outcome depends upon the provided image quality [25].

Pre-Upsampling Super Resolution

Image upsampling is the process of increasing the spatial resolution without changing the dimensional representation of the image. This principle safeguards the image during the zooming operation to avoid pixel breakage [26]. Examples of upsampling methods are traditional approaches such as bicubic interpolation, nearest-neighbour interpolation, bicubic spline interpolation, generalized bicubic interpolation, *etc.*

CHAPTER 6

Analyzing the Performances of Different ML Algorithms on the WBCD Dataset

Trupthi Muralidharr[1,*], **Prajwal Sethu Madhav**[1], **Priyanka Prashanth Kumar**[1] and **Harshawardhan Tiwari**[1]

[1] *Jyothy Institute of Technology, Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka India*

Abstract: Breast cancer is a disease with a high fatality rate each year. It is the most frequent cancer in women and the leading cause of death in women worldwide. The method of machine learning (ML) is an excellent way to categorize data, particularly in the medical industry. It is widely used for decision-making, categorization, and analysis. The main objective of this study is to analyze the performances of different ML algorithms on the WBCD dataset. In this paper, we analysed the performances of different ML algorithms, i.e., XGboost Classifier, KNN, Random Forest, and SVM (Support Vector Machine). Accuracy was used in the study to determine the performance. Experimental result shows that SVMs perform better and are more accurate than KNNs as the amount of training data increases. The SVM produces better results when the main component (PC) value grows and the accuracy rating exceeds the kNN.

Keywords: Breast Cancer, Decision Tree, Exploratory Data Analysis, Histograms, KNN, Random Forest, SVM, UCI Machine Learning Repository, WBCD, XgBoost.

INTRODUCTION

Breast Cancer is a critical illness seen mainly in women and has been one of the leading causes of their death. Breast cancer can mainly be detected by physical examination or techniques like Fine Needle Aspiration (FNA) and Mammography. These cancers start with a lump in the chest. Lumps formed in the chest can be benign or can be malignant. Benign tumors are the ones that are non-cancerous and malignant tumors are the cancerous tumors.

In 2020, there were 2.3 million women diagnosed with breast cancer and 685 000 deaths globally. As of the end of 2020, breast cancer was the world's most preval-

* **Corresponding author Trupthi Muralidharr:** Jyothy Institute Of Technology Pipeline Rd, near Ravi Shankar Guruji Ashram, Thathaguni, Karnataka India; E-mail: trupthimuralidharr@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

ent cancer [1]. Many women with breast cancer have no symptoms, so early detection helps the treatment process and increases survival rates. Machine Learning plays an important role in the early detection of breast cancer. Many algorithms are used in both supervised and unsupervised learning. The data set that is being implemented is the Wisconsin Breast Cancer Diagnosis (WBCD), and the methods being used are Decision Tree Classification using the Algorithms Regression tree, XG Boosting, Random Forest, and Ada-Boost [2].

LITERATURE REVIEW

Machine learning classifiers have long been used in medical research, particularly in breast cancer diagnosis. Classification is one of the most prevalent sorts of machine learning jobs. ML classifiers have been employed in various medical research initiatives, including the WBCD. According to previous research, the classifiers used in this study have a high level of classification accuracy [3]. Shokoufeh Alaei and Hadi Shahraki use the wrapper strategy. The study used three different breast cancer datasets and three different classifiers on Wisconsin breast cancer datasets: artificial neural network (ANN), PS-classifier, and genetic algorithm-based classifier (GA-classifier) (WBC). According to the findings, feature selection can aid classifiers in improving their accuracy, specificity, and sensitivity.

Jian Ping Li and his colleagues employed a recursive feature selection technique [4]. This method is used to train and test the classifier for the best prediction model. Additionally, performance evaluation criteria such as classification, specificity, sensitivity, Matthews' correlation coefficient, F1-score, and execution time have been used to assess the classifier's performance. The dataset "Wisconsin Diagnostic Breast Cancer" was used in this investigation. On this best subset of features, SVM achieved the best classification performance. Matthews' correlation coefficient is 99 percent, and the SVM kernel linear obtained high classification accuracy (99 percent), specificity (99 percent), and sensitivity (98 percent).

Li Chen and his colleagues suggest a combination of K-means and Boosted C5 (K-Boosted C5.0) based on undersampling to overcome the two-class unbalanced problem in breast cancer detection. To evaluate the new hybrid classifier's performance, it is applied to 12 small-scale and two large-scale datasets (Protein homology prediction and Breast Cancer) used in class unbalanced learning [5].

Their proposed hybrid approach outperforms most competing algorithms in terms of Matthews' correlation coefficient (MCC) and accuracy indices according to extensive experimental data.

The Quality Control Charts and Logistic Regression were used to diagnose breast cancer by Omar Graja, Muhammad Azam, and Nizar Bouguila. Data pre-processing, which involves removing outliers and reducing dimensionality to visualize the data, is a contribution of this study. Following the presentation of the data, the best machine learning algorithm for this diagnostic can be proposed. They employed the UCI machine learning breast cancer Wisconsin (original) dataset (WBCD) for the evaluation of our suggested approach [6].

The effect of dimensionality decreases utilization of free part examination (ICA) on bosom malignant growth choice emotionally supportive networks utilizing an assortment of classifiers, including fake neural organizations (ANN), k-closest neighbor (k-NN), spiral premise work neural organization (RBFNN), and backing vector machine (SVM) [7] were researched by Erdem Bilgili and Ahmet Mert. Thus, the characterization precision rates for thirty unique elements, except for RBFNN, have diminished hardly from 97.53 percent to 90.5 percent. Be that as it may, the RBFNN classifier is recognized with higher precision from 87.17 percent to 90.49 percent when utilizing the one-layered element vector.

Sulyman Age Abdulkareema and Zainab Olorunbukademi Abdulkareemb used the WBCD dataset to test the performance of two ensemble ML classifiers: Random Forest (RF) and eXtreme Gradient Boosting (XGBoost) [8]. The study's main goal was to assess the accuracy of the classifiers in terms of their efficiency and effectiveness in categorizing the dataset. This was accomplished by selecting all and reducing features from the dataset using the Recursive Feature Elimination (RFE) feature selection approach. According to the findings, XGBoost with five reduced features, and the RFE feature selection approach have the highest accuracy (99.02%) and the lowest error rate.

Liu et al. proposed a decision tree-based breast cancer prediction algorithm that balanced the training data using a sampling method. According to the experimental data, the proposed method generated good accuracy [9].

Krzysztof and Duch used the Separability of Split Value (SSV) criteria to create a heterogeneous forest of decision trees. The majority of their work involved extracting logical rules from diverse decision trees [10].

Stephan and Lucila contrasted logistic regression versus artificial neural networks in a range of medical categorization tasks. They proved the benefits of using logistic regression in the medical field from a technical aspect, as it is a white box model with interpretable results [11].

Salama and his colleagues [10] used the WEKA data mining technology to conduct a study titled "Breast Cancer Diagnosis on Three Different Datasets

CHAPTER 7

Application and Evaluation of Machine Learning Algorithms in Classifying Cardiotocography (CTG) Signals

Srishti Sakshi Sinha[1] and **Uma Vijayasundaram**[1,*]

[1] *M.Tech CSE, Pondicherry University, Puducherry, India*

Abstract: Cardiotocography (CTG) is a clinical procedure performed to monitor fetal health by recording uterine contractions and the fetal heart rate continuously. This procedure is carried out mainly in the third trimester of pregnancy. This work aims at proving the significance of upsampling the data using SMOTE (Synthetic Minority Oversampling Technique) in classifying the CTG traces. The project includes the comparison of different Machine Learning approaches, namely, Logistic Regression, Support Vector Machine (SVM), Naïve Bayes, Decision Tree, Random Forest, and K-nearest Neighbor (KNN) classifiers on the CTG dataset to classify the records into three classes: normal, suspicious and pathological. The results prove that applying SMOTE increases the performance of the classifiers.

Keywords: Classification algorithms, CTG, Decision Tree, Fetal Heart Rate, K-nearest Neighbors, Logistic Regression, Naïve Bayes, Random Forest, Support Vector Machine.

INTRODUCTION

Doctor Alan Bradfield, Orvan Hess, and Edward Hon invented Fetal Monitoring and Cardiotocography (CTG), a technical method of fetal monitoring by continuously recording the fetal heart rate (FHR) and the Uterine Contractions (UC) during pregnancy. The machine used to perform CTG is called Cardiotocograph, aka Electronic Fetal Monitoring (EFM). Later, Konrad Hammacher developed CTG for Hewlett-Packard.

CTG monitoring is used to assess fetal well-being and is performed in the third trimester of pregnancy. There are two methods for carrying out CTG: External and Internal. The external CTG can be used for continuous or intermittent monito-

* **Corresponding author Uma Vijayasundaram:** Assistant Professor, Department of Computer Science, Pondicherry University, Puducherry, India; Email: umabskr@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

ring. The FHR and UC are detected using the transducers attached to the mother's body.

The internal CTG is mainly used for more accurate tracking since it is detected by attaching electronic transducers to the fetal scalp. CTG is practiced because it assesses fetal health and well-being and identifies any changes that may or may not be associated with any complications or issues that may arise during pregnancy or labor. It is conducive to high-risk pregnancy conditions.

Different classification approaches have been applied for CTG analysis to date. Machine learning approaches like Support Vector Machine (SVM), Random Forest classifier (RF Classifier), K-nearest Neighbor (K-NN) [1, 2], decision tree, and Deep learning approaches like Artificial Neural Network (ANN), Deep Neural Network (DeepNN) [3], Convolutional Neural Network (CNN), Long Short-Term Memory Neural Network (LSTM-NN), *etc*. have been used till now. An advanced Deep Learning enhanced technology might be used to analyze the CTG traces in real-time with better accuracy than clinicians. This is why experiments and studies are going on in this field. This experiment studies different classification models in Machine learning (ML) and Deep Learning (DL) and their behavior on CTG datasets. Classification algorithms are used for predictive modeling where a class for a set of inputs is predicted. In simple terms, in machine learning, the data is used to train the algorithms, models are built based on the training, and decisions are made by the models based on what they have been taught. While deep learning algorithms learn and make intelligent decisions on their own. They have the capacity to execute feature engineering on their own. Deep Learning, a subset of Machine learning, is performed using advanced machine learning approaches. The features used in implementations make an important impact on how the algorithm will behave. Thus, feature engineering is a necessary procedure to be carried out while implementing any classification approach. Feature engineering mainly has two goals: to prepare an input dataset that is compatible with the learning models' requirements and improve the models' performance. Few of the techniques for feature engineering techniques are outlier detection, log transform, scaling, one-hot encoding, *etc*. The machine models used for classification in this study are Logistic Regression, the probabilistic classifier viz. Naïve Bayes classifier, Instance based classifier viz. K-nearest Neighbor, SVM, Decision Tree, and the ensemble classifier viz. Random Forest Classifier.

The following section provides a survey of the works related to CTG analysis. The architectural details follow this. The basic ML models are explained in the section Models and Methods. The experimentation details and result analysis are provided in further sections, along with performance measures and experimental analysis and results.

LITERATURE REVIEW

Since the invention of CTG, or continuous fetal heart record in 1957, many types of research have been carried out on and off since the mid-1980s. Different researchers have tried different approaches for the classification of CTG recordings. Czabanski *et al*. classified the fetal state using a hybrid model, viz. ANN-based on Logical Interpretation of Fuzzy rules [4]. Hakan Sakin *et al*. used a hybrid model comprising ANN and Simple Logistic for the same purpose [5]. Zafer Comert *et al*. published several papers in this field. They have done studies and experiments on different models like ANN and variations of Neural Network such as Deep CNN and other ML classifiers for CTG classification [6, 7], hypoxia detection [8], prediction, and assessment of the fetal distress during the trimester of pregnancy [9, 10]. Along with these, they have developed a useful software for CTG Analysis called CTG-OAS, which is an open-access software. The input can be a CTG recording in MATLAB file format. It offers several features like preprocessing of the CTG record and extraction of morphological, time domain, nonlinear and IBTF features, as well as application of SVM, ANN, and K-nearest Neighbor classifiers. Satish Chandra *et al*. published a paper on classification and feature selection approaches for CTG using ML techniques. They applied the same techniques that we have used but on a different dataset [1]. Noora Jamal *et al*. applied the Firefly algorithm and Naïve Bayesian classifier to the CTG dataset and achieved better prediction accuracy [11]. C. Ricciardi *et al*. have implemented SMOTE and various ML algorithms on the CTG dataset for the classification of the type of delivery from CTG signals [2]. Haijing Tang *et al*. and MS Iraji *et al*. have experimented on neural networks with different variations for the classification [12, 13].

Adulaziz Alsayyari experimented with Legendre Neural Network for CTG monitoring as well [14]. Bursa *et al*. have experimented with the use of CNN in biomedical data processing [15]. PA Warrik *et al*. classified intrapartum fetal state using an LSTM [16]. Paul Fergus *et al*. published a paper about their most recent work, which included the application of Multilayer Feedforward NN, one dimensional-CNN, and other ML classifiers for modeling segmented CTG time-series signals [17]. Zhidong Zhao *et al*. performed several experiments as well. Their recent investigation implemented 8- layer deep CNN and CWT on CTG signals for intelligent prediction of fetal academia [3]. This model achieved higher sensitivity (True Positive Rate) and specificity (True Negative Rate) when compared with other modern methods.

CHAPTER 8

Deep SLRT: The Development of Deep Learning based Multilingual and Multimodal Sign Language Recognition and Translation Framework

Natarajan Balasubramanian[1] and **Elakkiya Rajasekar**[1,*]

[1] *Research Scholar, Assistant Professor, School of Computing, SASTRA Deemed to be University, Thanjavur, Tamilnadu, India*

Abstract: Developing deep neural models for continuous recognition of sign gestures and generation of sign videos from spoken sentences is still challenging and requires much investigation in earlier studies. Although the recent approaches provide plausible solutions for these tasks, they still fail to perform well in handling continuous sentences and visual quality aspects. The recent advancements in deep learning techniques envisioned new milestones in handling such complex tasks and producing impressive results. This paper proposes novel approaches to develop a deep neural framework for recognizing multilingual sign datasets and multimodal sign gestures. In addition to that, the proposed model generates sign gesture videos from spoken sentences. In the first fold, it deals with the sign gesture recognition tasks using a hybrid CNN-LSTM algorithm. The second fold uses the hybrid NMT-GAN techniques to produce high-quality sign gesture videos. The proposed model has been evaluated using different quality metrics. We also compared the proposed model performance qualitatively using different benchmark sign language datasets. The proposed model achieves 98% classification accuracy and improved video quality in sign language recognition and video generation tasks.

Keywords: Convolutional Neural Network, Feature extraction, Generative Adversarial Networks, Long short-term Memory, Machine Translation, Recognition, Sign Language, Subunit modeling, Translation, Video generation.

INTRODUCTION

The World Health Organization (WHO) has reported five percent of the world's population, or approximately 466 million people globally, who have hearing problems. Sign Language (SL) plays a vital role in deaf and dumb communication. Sign languages are gesture-based and follow their own grammatical rules.

[*] **Corresponding author Elakkiya Rajasekar:** Research Scholar, Assistant Professor, School of Computing, SASTRA Deemed to be University, Thanjavur, Tamilnadu, India; E-mail: elakkiyaceg@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

Sign Languages are communicated through hand expressions and involve the eye, head, and mouth movements [1]. Ordinary people's recognition of sign gestures still poses enormous difficulties in understanding and replying. Due to these issues, communication between ordinary people and deaf-mute society is lacking. Although substantial research works are carried out in this area, the problems persist in handling continuous sentences. We investigate the recent advancements in deep learning techniques and propose a novel approach to construct a robust framework to address this issue. In addition to that, the different variants of sign languages added another difficulty in handling multilingual sign gestures and multimodal features. The proposed model is developed to address all these challenges as mentioned above. The proposed model follows the bidirectional approaches to recognition and video generation tasks. The recognition of sign gestures and translating them into text involves subunit modeling [2] and manual and non-manual feature extraction methods [3, 4]. The recognition accuracy of the model will be evaluated using generated text output and cost functions. On the other hand, the sign video generation from given spoken sentences is implemented using advancements in Generative Adversarial Network (GAN) models, Recurrent Neural Network (RNN) and its variants, and neural machine translation techniques [5].

The sign video generation approach involves the translation of spoken sentences into sign glosses and mapping them with sign gesture images. The frame sequences will be arranged properly and produce a plausible video output. In general sign language, hand gestures are divided into four categories: manipulative, conversational, controlling, and communicative. Sign Language is based on the communicative type of gestures considered testbed for computer vision techniques [6]. Two main approaches are sensor-based and vision-based approaches. Sensor-based techniques use electronic devices such as surface Electromyography (sEMG)] [7], ECG [8], Data gloves [9], colored gloves [10], depth cameras [11], Kinect Sensors [12] and Leap Motion Sensors [13]. Wearing these devices creates discomfort for normal people and lacks a user-friendly approach. Vision-based techniques employ deep artificial models to recognize images or videos directly, explore deep insights about gesticulation information and benefit the users. Image segmentation, classification, and feature extraction techniques are employed for continuous DSL recognition. The classifiers such as Hidden Markov Models (HMM) [14 - 17], K-Nearest Neighbour (KNN) [18], SVM - Support Vector Machine [18], Convolutional Neural Network (CNN) [19 - 24], RNN [25], LSTM - Long short term memory [26], Recurrent CNN [27], and Gated Recurrent Unit (GRU) [28] Techniques attained wider attention in Sign Language recognition and Translation (SLRT) research. Many research works have been carried out in past decades to improve Sign Language communication using smart devices. SL entails the human body's various segments, specifically fingers, hands, mouth, arm, head, body movements, facial postures, and gaze movements to deliver information.

The recognition of continuous gestures of signers is challenging because of the occlusion, articulation, and epenthesis movements of sign language. The proposed model aims to recognize the signs accurately in a real-time, continuous, and automated way. The various dimensions of research in sign language communication are elaborated (Fig. **1**).

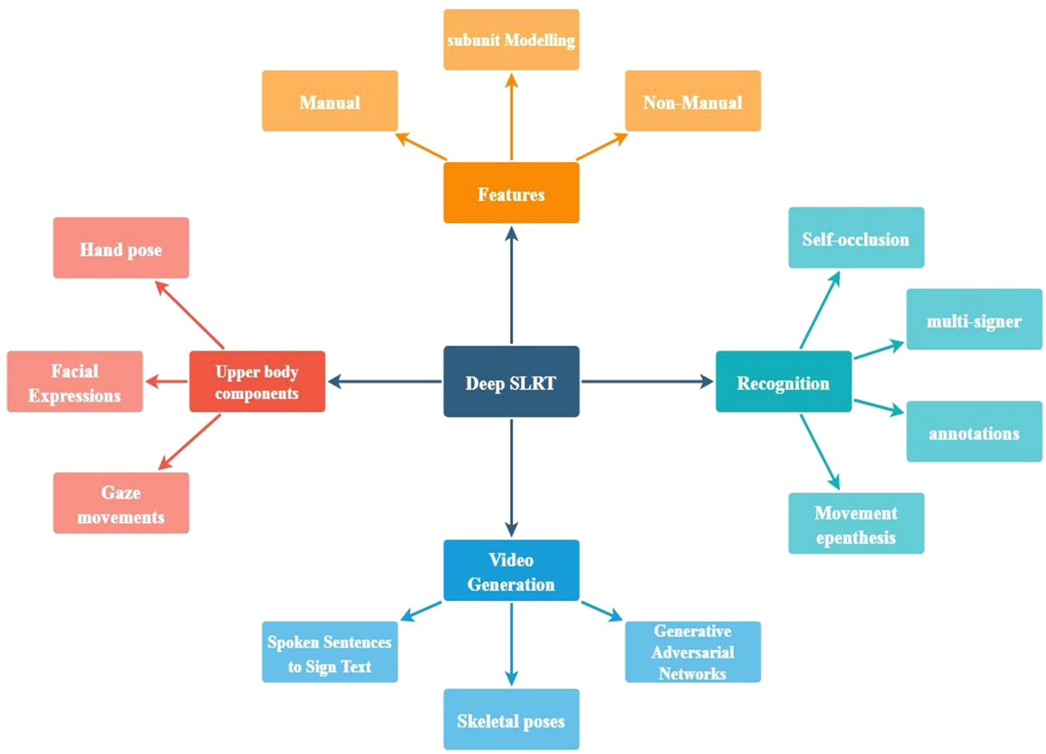


Fig. (1). Various dimensions of Sign Language based Deep learning Research.

The objectives of this work are listed below:

- Development of continuous sign corpus for Indian Sign Language (ISL) using multi-signers and various background conditions and lighting. This corpus mainly aimed to cover 100 common English words used daily.
- Development of SLRT Model for ISL recognition, translation, and sign video

CHAPTER 9

Hybrid Convolutional Recurrent Neural Network for Isolated Indian Sign Language Recognition

Elakkiya Rajasekar[1], **Archana Mathiazhagan**[1] and **Elakkiya Rajalakshmi**[1,*]

[1] *School of Computing, SASTRA Deemed University, Thanjavur 613401, India*

Abstract: Even though the hearing and vocally impaired populace rely entirely on Sign Language (SL) as a way of communication, the majority of the worldwide people are unable to interpret it. This creates a significant language barrier between these two categories. The need for developing Sign Language Recognition (SLR) systems has arisen as a result of the communication breakdown between the deaf-mute and the general populace. This paper proposes a Hybrid Convolutional Recurrent Neural Network-based (H-CRNN) framework for Isolated Indian Sign Language recognition. The proposed framework is divided into two modules: the Feature Extraction module and the Sign Model Recognition module. The Feature Extraction module exploits the Convolutional Neural Network-based framework, and the Model recognition exploits the LSTM/GRU-based framework for Indian sign representation of English Alphabets and numbers. The proposed models are evaluated using a newly created Isolated Sign dataset called ISLAN, the first multi-signer Indian Sign Language representation for English Alphabets and Numbers. The performance evaluation with the other state-o--the-art neural network models have shown that the proposed H-CRNN model has better accuracy.

Keywords: Hybrid Neural Networks, Isolated Indian Sign Language, Image segmentation, Sign Language Recognition.

INTRODUCTION

Sign Language is a medium of communication for the deaf and mute community. As per the World Health Organization, approximately 6.1 percent of the worldwide population is deaf or with speech impediment. According to the National Association of the Deaf (NAD), about 18 million people in India suffer from hearing impediments. Sign language differs from region to region. There are approximately more than 120 Sign Languages (SL). Some of the notorious sign languages include Arabic [1], American, Bhutanese [2], German, Chinese [3], Russian, Turkish [4], and Indian Sign languages.

* **Corresponding author Elakkiya Rajalakshmi:** School of Computing, SASTRA Deemed University, Thanjavur, India; E-mail: elakkiya@cse.sastra.edu

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

The Indian Sign language itself differs from region to region. Deaf and mute people generally need someone to translate their sign language for their regular communication. The formulation of the Sign Language Recognition System (SLR) is motivated by the desire to contribute to the development of feasible solutions for the automation of sign language recognition, hence eliminating the need for human assistance during interpretation. There have been many recent developments in building SLR systems. Most of the state-of-the-art SLR systems that are reliable, accurate, and simple are dependent on the wearable devices such as gloves. Although wearable devices give accurate predictions based on the measures of the finger joints, the orientation of the hand and the location of key points on the face, *etc.,* when it comes to practical usage for day to day life of a person with hearing and speech impediment it would be tedious, restrictive, bizarre as well as costly to use. The robustness of an image/video-based SLR system's image processing framework, *i.e.,* its potential to withstand heterogeneous, dynamic, or generally unrestrained environments, poor lighting, and different skin tones, has a significant impact on its usability. One of the other main drawbacks in building the most accurate model for SLR is the lack of publicly available large datasets. There are no publicly available datasets for Indian Sign Language (ISL). Though the state-of-the-art models have used their customized datasets, they have not considered the dataset that is signer-independent, having different illuminations and backgrounds.

To overcome the above-said issues, we have proposed an image/video-based SLR system using a Convolutional Recurrent Neural Network (CRNN) framework to comprehend sign language by feeding it images/videos of hand gestures in different illumination scenarios with different backgrounds, different hand orientations, and different signers. We have created our own dataset, namely ISLAN [5], which consists of Indian gestures for Alphabets and numbers, and it is made publicly available for further research. The proposed CRNN framework can be divided into two modules, namely Feature Extraction (FE module) and Model Recognition (MR module). The FE module has been implemented using Convolutional Neural Network (CNN), and the MR module has been implemented using Recurrent Neural Network (RNN). By integrating both, we get an H-DNN, namely CRNN model. The main objective of the proposed work is:

a. Creation and collection of Indian Isolated Sign Language dataset for Indian Alphabet and Numeric signs (ISLAN) and making it publicly available.
b. Implementation of the H-CRNN SLR model for feature extraction and sign language recognition.
c. Evaluation of the model proposed using the newly created ISLAN dataset.

The rest of this paper is organized as follows: Section 2 gives a brief discussion of related works and a literature review. Section 3 introduces the proposed methodology, followed by result analysis and discussion in section 4. Section 5 discusses the conclusion and future work.

RELATED WORK

During the last couple of years, studies in the area of sign language gesture recognition have made tremendous progress. The present findings serve as a foundation for future applications intended to assist deaf people in their assimilation into hearing society. There has been much research conducted on the Sign Language Recognition System. There have been many sensor-based as well as video-based recognition systems used for formulating the SLR system. Many researchers have worked on traditional methods without the implementation of neural networks and machine learning [6]. One of the approaches proposed English alphabet recognition using scale-invariant feature transform [7]. A framework for recognition of the Indian Alphabet gesture was proposed, which dealt with the Eigenvalues and Eigenvectors [8]. Various HMM-based models were also introduced for recognition of Arabic SLR [9 - 11]. Then came the era of sensor-based gesture recognition using sensors such as Leap motion sensors and Microsoft Kinect [12], other machine learning approaches such as KNN, SVM, *etc* [13]. Some approaches also used more than one digital camera along with wearable sensors to capture the gestures from different angles [14]. There were several other frameworks introduced having wearable sensors with the integration of Neural Networks. For instance, a CNN-based framework was proposed using a Microsoft Kinect sensor using American Sign Language consisting of an alphanumeric dataset [15]. The same combination of the framework was also utilized using Italian Sign Language, where the framework was modeled based on the hand and upper body features [16]. Later, a neural network model was proposed for Arabic SLR where not only the height and width dimension but the depth dimension of the object was also extracted from the sensor and fed into various CNN frameworks [17]. Another Arabic SLR framework was proposed, which utilized a comparatively large dataset consisting of 7869 signs for Alphabets and Numbers, which yielded a good accuracy of 90% [18].

Many kinds of research have been out carried on sign language recognition using the video-based approach in different regions throughout the world. The researchers have exploited various CNN-based frameworks [19], such as VGG16, VGG19, AlexNet, ResNet, inception net, *etc.,* for sign language classification. A Chinese SLR system was proposed that integrated CNN with Long-Term Memory Network (LSTM) [20] for Isolated SLR using word sign videos which yielded a 95% accuracy level [21]. The CNN-based frameworks have a major drawback

CHAPTER 10

A Proposal of an Android Mobile Application for Senior Citizen Community with Multi-lingual Sentiment Analysis Chatbot

Harshee Pitroda[1,*], **Manisha Tiwari**[1] and **Ishani Saha**[1]

[1] *Department of Computer Engineering, NMIMS University, Mukesh Patel School of Technology Management & Engineering, Mumbai, India*

Abstract: Throughout these years, technology has transformed our world and has become an integral part of our everyday lives. This massive digitalization needs to take into consideration the elderly community too. However, the elderly community is usually digitally excluded, but despite these various and diverse difficulties, they remain driven, engaged, and eager to make an effort, to incorporate developing digital technologies into their daily life. Hence, this research implements the functionality of analyzing and determining the emotions of senior citizens by utilizing various natural languages processing-based machine learning techniques like SVM, Random Forest, and Decision Tree. Different algorithms were used, and various parameters were compared to obtain the results. It is seen that SVM gave the best results with an accuracy of 88%. The android application "Bandhu – your forever friend" would be accessible and usable by the elderly community. In addition, this proposed application offers various other features, including learning a new language and listening to the regional songs which would cater to the multicultural requirements in a diverse country like India, recording their audio stories to preserve the ethnic culture and inspire others, and lastly becoming tech-savvy by following easy video tutorials. These features would not only keep them engaged but also tackle their loneliness and isolation to some extent, as all these features were considered after surveying around 100 senior citizens' lifestyles and needs. This app will also make them more digitally independent. All the above-mentioned functionalities have been implemented using the programming language Java and the android application is built in Android Studio. Also, the entire app is in the Hindi language, considering that this language is the most preferred and spoken language in India.

Keywords: Android application for senior citizens, Decision Tree, Multi-lingual sentiment analysis chatbot, Random Forest, Support Vector Machine (SVM), TF-IDF Vectorization.

* **Corresponding author Harshee Pitroda:** Department of Computer Engineering, NMIMS University, Mukesh Patel School of Technology Management & Engineering, Mumbai, India; E-mail:harshee.pitroda2910@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

INTRODUCTION

The world is changing at a breakneck pace, and although much emphasis is placed on developing the potential of the youth community, the other end of the age spectrum, that is, the elderly community, is also increasing at a rapid pace often overlooked. Life expectancy is anticipated to rise further in the next decades as medical science advances, signaling an urgent need to build a digital infrastructure that supports and leverages the potential of this enormous but often-forgotten 'silver' sector. These technological advancements should seek to improve the lives of the elderly community by making them easier and more enjoyable, hence helping them to cope with the feeling of loneliness, isolation, and independence. So, the app "Bandhu – your forever friend" aims to cater to the elderly community of society.

The proposed solution is built incrementally, augmenting various features. After each successive version of this application, another functionality was added to the preceding version until all intended functionalities were implemented.

A survey was conducted through google forms, where around 100 senior citizens were asked a series of questions, and their responses were recorded. The questions asked are listed in Section VI, named Questionnaire. The survey results were used as a guiding tool to decide on various features to be implemented in this proposed solution.

Certain physiological and cognitive changes are nearly inevitable as people age, and these changes must be incorporated while designing the app's user interface. In designing the user interface of the app, all these factors are taken into consideration.

The salient features of this app's UI design include the following:

- The entire application is in the Hindi language, considering that it is the most preferred and spoken language in India.
- Hand-eye coordination may decline as people age, which may impair UI interaction. As a result, in this app, all UI components that need to be read or clicked are made larger in size.
- As motor abilities may deteriorate with age, it becomes more difficult for the elderly to operate a smartphone in many ways. So, this app's user interface is made incredibly simple to use and navigate.
- Colour vision may also deteriorate with aging, making it harder to distinguish between different hues. As a result, contrast colors have been utilized to help the senior citizens clearly identify between the background and the buttons or texts.

- Icons are labeled with text everywhere, making the function of the iconography clear to them.

The modules of the app “Bandhu – your forever friend” are as follows:

- **Multi-lingual sentiment analysis chatbot** - Sentiment analysis is the process of analyzing and determining the emotion or purpose behind a piece of text, voice, or any other method of communication. This is performed by using the Natural Language Processing-based Machine Learning model. The goal of conducting a sentiment analysis is to identify the sentiments underlying the senior citizen's texts and then respond to them accordingly. This chatbot can assist fill the void of an elderly population in need of human connection, as well as alleviate feelings of loneliness and isolation. It can also be useful in aiding senior citizens, both as a companion and as an instrument to keep their physical and mental health in check.
- **Listen and learn a language** - learning a new language at a later age may help one to refresh and invigorate both social and cognitive skills. Learning a new language exercises the brain and has been shown to improve IQ, memory, and attention. This app also allows the elderly community to learn by listening to the language, and also at their own speed, without the aid of youngsters.
- **Learn technology** – Learning new technology will ease their life, make them more independent, and they won't remain secluded from society. This is achieved by providing easy video tutorials in the Hindi language to show them to do online shopping, send and receive emails, pay bills, do banking services, learn how to utilize social media, download and upload materials, *etc.* They can learn these at their own speed without the aid of the young.
- **Audio Diaries** – Seniors, who reflect on their lives, are typically astonished and grateful for how much they've been through. Having someone to share one's life tales boosts the senior citizen's self-esteem. It is a fantastic legacy to preserve life tales through audio recordings. Through this app, senior citizens can record their own cultural stories, recipes, and real-life inspiring and motivating experiences, to share with others.
- **Listen to regional songs** - Music is a powerful tool for evoking memories and increasing people's overall quality of life. Senior citizens would enjoy listening to regional music from this app, which would recall their wonderful memories.

The remainder of this paper discusses in detail the multi-lingual sentiment analysis chatbot module and is organized as follows: Section II is the literature review. Section III is further subdivided into three subsections, A, B, and C, describe the proposed framework for the multi-lingual sentiment analysis chatbot, proposed architecture for integrating the model, and the dataset, respectively.

CHAPTER 11

Technology Inspired-Elaborative Education Model (TI-EEM): A futuristic need for a Sustainable Education Ecosystem

Anil Verma[1], **Aman Singh**[1,*], **Divya Anand**[1] and **Rishika Vij**[2]

[1] *Department of Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India*

[2] *Department of Veterinary Physiology & Biochemistry, Dr. GC Negi College of Veterinary & Animal Science, Palampur, Himachal Pradesh, India*

Abstract: Before three decades, providing higher education infrastructure for young aspirants in their locality was a challenge for India. With 5164 higher education institutions, including universities, colleges, and stand-alone institutions, India has surpassed the United States as the global leader in educational infrastructure over the last two decades. This work intends to propose an elaborative education ecosystem for sustainable quality education. The secondary data from top global ranking agencies (Times, QS, Webometric, Scimago, and Shanghai Ranking) is deployed to avoid the cost of a worldwide survey for primary data and the execution time. Quality education's quantitative and qualitative parameters are reviewed separately on different scales. The need for the proposed model is evaluated on academic reputation, employer reputation, faculty-student ratio, citations per faculty, international faculty, international students, and infrastructure on the 7-point quality scale. The proposed elaborative model will establish a robust quality education ecosystem on global parameters. The proposed model emphasizes the use of emerging technologies including the Internet of Things (IoT), Artificial Intelligence (AI), and Blockchain (BC), in the education industry.

Keywords: Education Ecosystem, Higher Education, Quality Education, Technology-Enabled Infrastructure.

INTRODUCTION

More than 993 universities, 39931 colleges, and 10725 stand-alone institutions by 2019 [1] made India the top country in the world regarding the number of universities and colleges.

* **Corresponding author Aman Singh:** Department of Computer Science and Engineering, Lovely Professional University, Jalandhar, Punjab, India; E-mail: amansingh.x@gmail.com.

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

As per webometric data July-2019, India is the top participant with 3944 colleges and universities in the webometric university ranking survey 2019 [2]. The United States of America was just behind India with 3257 participants. China is the most populated country in the world at number 3, with 2208 participants. Indonesia, Brazil, and the Russian Federation are the countries with more than 1000 participants each. A total of 28077 colleges and universities participated in this survey. The top six countries, i.e., India, the US, China, Indonesia, Brazil, and Russia, cover almost 50% of the total number of colleges and universities around the globe. Two hundred-eleven countries across the world participated in this survey in 2019. Only India covers 7.12% of the world's population of colleges and universities in this survey. China has only 0.1 billion more inhabitants than India, but India has 1736 more colleges and universities than China (and 687 more colleges and universities than the US). So, India is nowhere lagging in terms of infrastructure facilities as India has more colleges and universities compared with any country in the world.

Infrastructure is only one aspect of the education system; intellectual faculty and a suitable environment for learning are the other two aspects that also have equal significance. The QS-World University Ranking—2020 justifies the intellect of the Indian faculty. The appreciable research work conducted by Indian faculty on a global platform in 2019 ranked India among the top three nations in the world in citation per faculty [3]. More than 40 domain-specific apex bodies, including UGC, AICTE, PCI, MCI, BCI, AISHE, NIRF, and NBA, are deployed in India to ensure a better learning environment at the ground level. Although all these apex bodies have a rich technological environment for quality assurance, the position of quality education in India on a global platform is even below the world's average quality. In India, the basic education model that comprises infrastructure, faculty, and a learning environment is deficient in achieving sustainable quality education. India has already completed most of the quantitative goals in the higher education sector, but the qualitative goals are not meeting the average standards on global platforms. Fig. (**1**) illustrates an elaborative model with many other factors contributing to quality education. All the inner elements required more focus on quality enhancement. The exploration and systematic assessment of all inner elements (including curriculum, activities, exposure for students and faculty, awareness for quality education, participation in quality assessment, faculty interest in imparting quality education along with their academic and research growth, students' interest and urge for quality education, communication and dialogue among all stakeholders, including students, teachers, and respective apex bodies), will edify a comprehensive model for sustainable quality education [4].

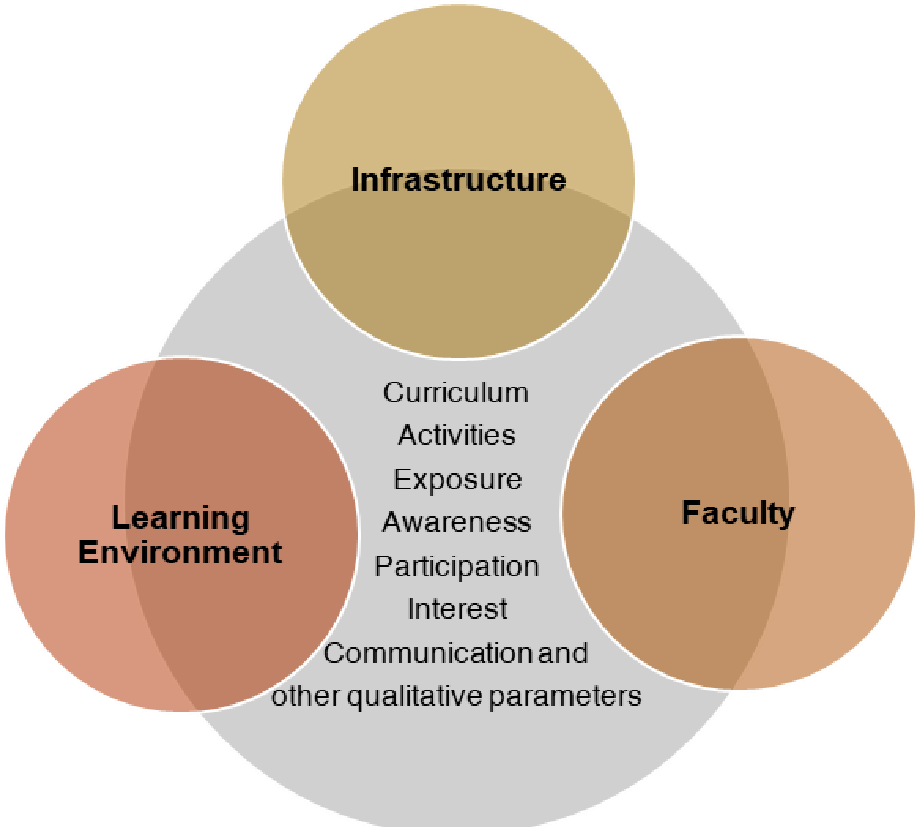


Fig. (1). Elaborative Education model with various qualitative parameters.

This study reviews the global higher education system using quantitative and qualitative parameters. The comparative measures used in this study are qualitative parameters with those of the top 5 countries in education and quantitative parameters with those of the top 5 most populated countries in the world. Moreover, figuring out the futuristic quantitative and qualitative needs for sustainable quality education in India is the central aim of this study, which further proposes a technology-inspired solution for sustainable quality education.

The rest of the paper is organized as follows. Section II discusses the background of the Indian education system. Section III presents the methodology and objectives of this study. Section IV illustrates a detailed discussion of a quantitative and qualitative comparison of various effects with their respective results. Section V concludes the paper.

CHAPTER 12

Knowledge Graphs for Explaination of Black-Box Recommender System

Mayank Gupta[1] and **Poonam Saini**[1,*]

[1] *Department of Computer Science and Engineering, Punjab Engineering College, Chandigarh, India*

Abstract: Machine learning models, particularly black-box, make powerful decisions and recommendations. However, these models lack transparency and hence cannot be explained directly. The respective decisions need explanation with the help of techniques to gain users' trust and ensure the correct interpretation of a particular recommendation. Nowadays, Knowledge graphs (*K*-graph) has been recognized as a powerful tool to generate explanations for the predictions or decisions of black-box models. The explainability of the machine learning models enhances transparency between the user and the model. Further, this could result in better decision support systems, improvised recommender systems, and optimal predictive models. Unfortunately, while these black box devices have no detail on the reasons behind their forecasts, they lack clarity. White box structures, on the other hand, will quickly produce interpretations due to their existence. The chapter presents an exhaustive review and step-by-step description for using knowledge graphs in generating explanations for black-box recommender systems, which further helps in generating more persuasive and personalized explanations for the recommended items. We also implement a case study on the MovieLens dataset and WikiData using *K*-graph to generate accurate explanations.

Keywords: Black-Box Recommender System, Explainability, Knowledge Graph.

INTRODUCTION

Machine learning models, essentially the black-box models, generate decisions and recommendations; however, lack of transparency and direct explainability make them untrustworthy to end-users. Therefore, such decisions need explanation using techniques that are interpretable as well as accurate to make any specific recommendation. The high-quality explanations of such recommendations can boost trust, effectiveness, efficiency, and satisfaction [1].

* **Corresponding author Poonam Saini:** Department of Computer Science and Engineering, Punjab Engineering College, Chandigarh, India; E-mail: poonamsaini@pec.edu.in

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

Introduction to Recommender System

Recommender systems have been gaining attention in recent times as they can ease information overload. One of the major concerns is the explainability of recommender systems. The authors [2] mentioned that fairness issues had received a lot of publicity lately, particularly in the sense of intelligent decision-making processes. However, the explainable feedback frameworks, for example, can be susceptible to both explanation prejudice and success disparity.

Owing to the design of collective filtering, inactive users may be more vulnerable to getting unsatisfactory feedback due to a lack of training evidence, and their recommendations may be affected by the training records of active users, resulting in unequal treatment by the framework. Recommender systems are constantly being used to forecast consumer needs on websites and to suggest specific solutions that will help them deal with knowledge overload. In recommendation systems, existing model-based collective filtering algorithms, like latent factor models, are considered state-of-the-art [2]. Fig. (**1**) shows the google trend for research interest in explainable artificial intelligence (XAI) key terms over time.

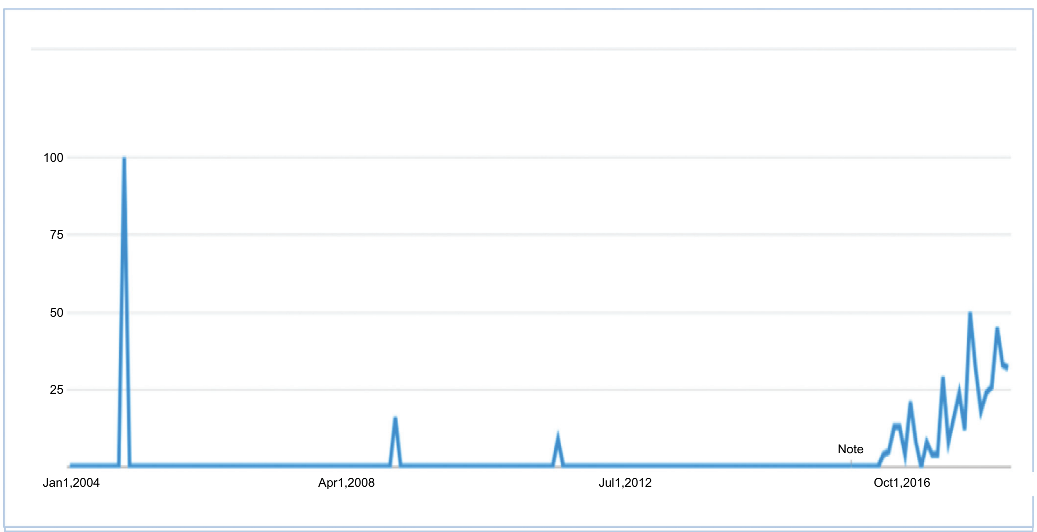


Fig. (1). Google trends result for research interest of Explainable Artificial Intelligence term [3].

Introduction to Knowledge Graphs

A Knowledge Graph (KG) could be used to generate explanations for the predictions or decisions made by the black box models. The KGs are a modern way of data representation that would find a lot of usage in generating explanations in the coming future. This could be useful in making better DSS (Decision Support Systems), Recommender Systems, Predictive Models, and so on. KGs usually provide much more valuable data and links regarding objects than other forms of side information. KG is a guided heterogeneous graph in which the nodes represent individuals, and the edges represent relationships. Several academic KGs, such as NELL and DBpedia, as well as commercial KGs, such as Google Knowledge Graph and Microsoft Satori, have recently been proposed. Many implementations of these information graphs have been good, including KG completion [4], query answering [5], word embedding [6], and text classification [7]. Researchers tried to use KGs to boost the efficiency of recommender mechanisms after observing their effectiveness in diverse domains in handling a variety of problems.

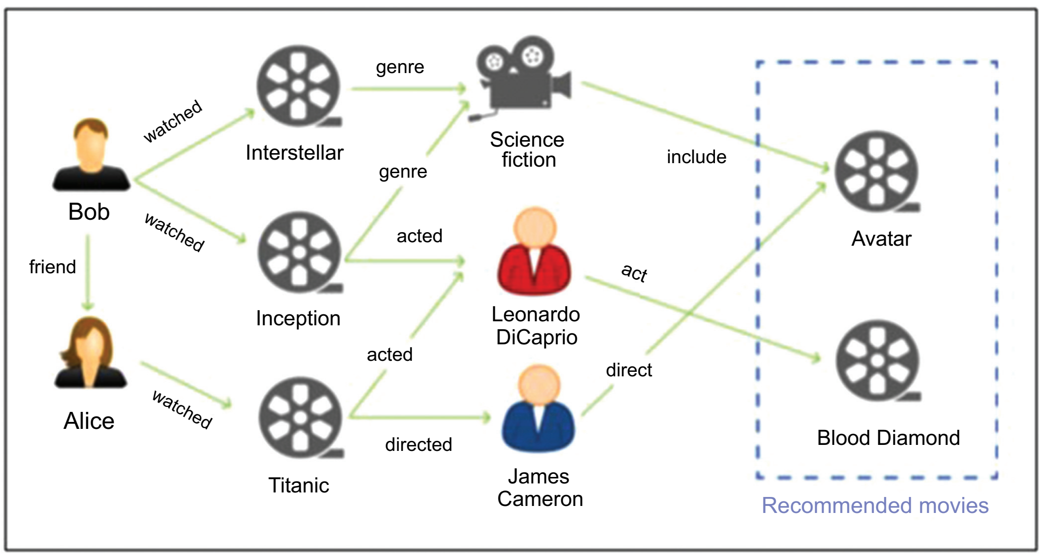


Fig. (2). An illustration of KG-based recommendation [8].

Fig. (**2**) provides an example of KG-based recommendation, in which Bob is recommended for the films “Avatar” and “Blood Diamond.” The users, plays, writers, directors, and genres are all individuals in this KG, with interactions, belonging, performing, controlling, and fellowship serving as connections between them. The movies and users are associated with various latent connections using KG, which ensure recommendation accuracy. The authors [9] show that the explainability of referral outcomes is an advantage for such schemes. Following the connection sequences in the user-item graph will show the reasons for recommending these two movies to Bob in the same case. For

CHAPTER 13

Universal Price Tag Reader for Retail Supermarket

Jay Prajapati[1,*] and **Siba Panda**[1]

[1] *Department of Data Science, SVKM's NMIMS, Mumbai, Maharashtra, India*

Abstract: Retail supermarkets are an essential part of today's economy, and managing them is a tedious task. One of the major problems faced by supermarkets today is to keep track of the items available on the racks. Currently, the track of the product on the shelf is kept by price tag readers, which work on a barcode detection methodology that has to be customized for each store. On the other hand, if barcodes are not present on the price tags, the data is manually fed by the staff of the store, which is really time-consuming. This paper presents a universal pipeline that is based on Optical Character recognition and can be used across all kinds of price tags, and is not dependent on barcodes or any particular type of price tag. This project uses various image-possessing techniques to determine and crop the Area of Interest. It detects the price of the product and the name of the product by filtering the OCR outputs based on the area and dimensions of the bounding boxes of the text detected. Additionally, the presented pipeline is also capable of capturing discounted prices, if any, for the products. It has been tested over price tags of five different types, and the accuracy ranges from 78% to 94.5%.

Keywords: Artificial intelligence, Contour detection, Image processing, Optical character recognition, Retail supermarkets.

INTRODUCTION

Currently, the price tag reading is widely done by a simple barcode scanner, in which the barcode is scanned, the code is matched with the database, and the information is extracted. The issue with this system is the anomaly of the barcode from store to store. Each of the stores may have its own barcode according to its internal inventory management systems. So the price tag reader has to be custom-tailored for each store. It is quite difficult to maintain data about each store. Also, some stores might not have these kinds of barcode-based inventory management systems, and the price tags do not have barcodes. This is a reason why such syste-

* **Corresponding author Jay Prajapati:** Department of Data Science, SVKM's NMIMS University, Mumbai, Maharashtra, India; E-mail: jayprajapati1141@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

ms are not used often. So this project will be using a very popular technique called optical character recognition which is widely used nowadays in number plate detection and recognition, which will be able to read the name, price, and discounted price(if any) of the product irrespective of the type of the tag.

LITERATURE REVIEW

There has been research work done on this front, and the closest one is- Price Tag Recognition using HSV Color Space [1]. This system requires a format and color of the price tag beforehand. It detects the price tag from the image using HSV color filtering and finds the largest rectangle contour in the image. The image is then cropped according to the largest contour, giving the cropped image the price tag. Then the price tag image is masked with a predefined mask image for that particular type of price tag (the research work has considered 3 predefined price tag formats- A, G, T). It returns two cropped images- one is the price-cropped image, and the second one is the product name image. Finally, an OCR is run over these two images to read the price and product name.

The problem with this system is that it requires the format and color of the price tag beforehand, and there might be n number of price tag designs, and there might be n number of colors for price tags, so the system is required to maintain the mask images for each format. Secondly, it doesn't consider the discounts and offer prices for products, and it only reads the price and name of the product.

Computer Vision and Image Processing: A Paper Review' [2] provides a deep, comprehensive review of many computer vision and image processing techniques in the past. It talks about the image processing applied to digital sound systems, food analytics, and CNN used for object detection. It also talks about recent pattern recognition research and image segmentation.

'Two-stage approach to extracting visual objects from paper documents' [3] is a two-stage approach for detecting components like logos, signatures, text blocks, tables, and stamps from the documents. The first stage of the approach is a rough detection through a simple and fast approach called Adaboost Cascade for the classifier, and the second stage is verifying stage, which uses additional techniques like First-order statistics (FOS), Gray-level run-length statistics (GLRLS), Haralick's statistics (HS), Neighboring gray-level dependence statistics (NGLDS), Low-level features (LLF), Histograms of oriented gradients (HOG) and local binary patterns (LBP). It was found that the accuracy was not that great in the first stage, but it improved after the verification stage, except for the signature, in which the accuracy decreased after the verification stage.

'Malaysian Car Number Plate Detection System Based on Template Matching and Colour Information' [4] this research paper proposes a pipeline for detecting number plates. This approach goes about cropping the lower half of the image, continued by pre-processing the image by using top fat filtering, thresholding, intensity correction, and median filtering. Then to identify the region of the number plate, template matching is used, pixel-by-pixel matching is done, and a numeric index is calculated, which shows how good the matching performance is. Then to reduce the noise in the image, contrast correction is used. To tackle the number plates with different dimensions, the bounding box is made with a padding of 10 pixels. The outputs are good for the number plates with similar dimensions as the masking template but don't perform well for larger number plates and number plates in different colors' Post-OCR parsing: building simple and robust parser *via* BIO tagging' [5]; this research talks about a downstream task of parsing textual information embedded in images. It has taken into account parsing tasks for name cards and receipts. There are four major steps in this process- OCR, Serialization, BIO tagging, and parse generation. It uses in-house OCR, which consists of a CRAFT text detector and Comb. Best text recognizer. The OCR outputs are serialized by sorting them from top to bottom and left to right according to their coordinates. These serialized tags are then BIO parsed using a neural network. Input vectors are mapped by adding coordinates, tokens, sequential positions, and segments. The model gives outputs of mixed BIO tags, which can be decoded according to the target field. These outputs are further refined. This approach enables information propagation in abstract space and removes the need for customization.

'A Novel Method for Indian Vehicle Registration Number Plate Detection and Recognition using Image Processing Techniques' [6] this research paper presents a pipeline for detecting and reading license plates. The approach is divided into three major parts- pre-processing, detection and recognition. The pre-processing includes image resizing, RGB to HSV conversion, greyscale extraction, morphological transformations, Gaussian smoothing, and Inverted Adaptive Gaussian Thresholding. The second step includes applying contours to the image, followed by filtering and grouping of contours, correcting the angle of the license plate, and finally fixing overlapping characters and contours. Once the license plate is detected, a model is trained using the KNN algorithm to recognize the characters and digits of the license plate. This research helped tackle situations of blurred images, images with varying illumination, noisy images, and non-standard license plates.

'Text Extraction from Bills and Invoices' [7], this research paper has presented a pipeline for extracting text from bills and invoices. This approach uses OpenCV and Tesseract OCR. The approach goes about by detecting and cropping the piece

CHAPTER 14

The Value Alignment Problem: Building Ethically Aligned Machines

Sukrati Chaturvedi[1,*], **Chellapilla Vasantha Lakshmi**[1] and **Patvardhan Chellapilla**[1]

[1] *Dayalbagh Educational Institute, Agra, India*

Abstract: Autonomous systems are increasingly being employed in almost every possible field. Their level of autonomy in decision-making is also increasing along with their complexity leading to systems that will soon be making decisions of utmost importance without any human intervention at all or with the least human involvement. It is imperative, therefore, that these machines be designed to be ethically aligned with human values to ensure that they do not inadvertently cause any harm. In this work, an attempt is made to discuss the salient approaches and issues, and challenges in building ethically aligned machines. An approach inspired by traditional Eastern thought and wisdom is also presented.

Keywords: Artificial Intelligence, AI system, Machine ethics, Values, The Eastern perspective of intelligence.

INTRODUCTION

Autonomous systems are being developed for performing tasks in different fields. These systems have embedded AI, which gives them the power of decision making. For instance, Deep Blue [1], is an AI-based chess game that has beaten the world chess champion. It all started with Artificial Narrow Intelligence, where the systems are capable of performing tasks such as playing games like Atari [2] and Go [3, 4]. ANI is the first level of AI, and an ANI system specializes in any one area only. The appearance of this level of AI development is clearly visible in the technologies that we use every day. These advancements will increase with every coming year, and technology will become more complex and pervasive as it becomes more intelligent. Artificial General Intelligence is the next level of Artificial Intelligence. At this level, autonomous systems can perform any intellectual task that a human can perform, meaning it has the ability to solve the

* **Corresponding author Sukrati Chaturvedi:** Dayalbagh Educational Institute, Dayalbagh, Agra, India; E-mail: sukratichaturvedi@dei.ac.in

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

problem, learn from experiences, reason, *etc.* The third and the last level is Artificial Super Intelligence, which will surpass human intelligence and be practically much smarter than the best human brain available in every field [5].

Currently, the lowest level of AI, i.e., ANI, has been achieved by humans in many domains. Examples include Google search, email spam filters, *etc.* The current effort that has heralded the AI revolution is the path from ANI, through AGI, to ASI – a path that will change everything around us. These advancements in AI appear to indicate that AGI is achievable in the near future, i.e., by the next five to ten years, although there are some who are more conservative in their predictions.

If existing AI systems misbehave, they can be easily monitored and shut down or modified. Examples where problems in data caused the AI system to behave unacceptably, are as follows:

- Google's photos classification algorithm tagged dark-skinned people as gorillas because it has not been trained with enough examples of people with dark skin.
- In 2018, Amazon had to shut down a machine learning tool used to make hiring decisions because it exhibited bias against women. This was also because the data it was trained with had this bias.

However, if a super-intelligent AI system misbehaves, it would realize that modifying or shutting down might interfere with its capability to accomplish its goals. If the super-intelligent AI system, therefore, decides to resist shutdown and modification, it would be smart enough to surpass its programmers if the programmers have taken no prior precautions. The issue of building AI systems that will help their developers in accomplishing the task build systems that would not inadvertently harm their developers, as well as the society, is known as the AI control problem. One major concern is that we have to solve the AI control problem before developing a super-intelligent AI system.

A major approach to solving the AI control problem is to include alignment into the system- the aim here is to align the AI system with human values so that it does not harm humans or gain control over humans.

Value Alignment Problem

Progress in creating AI systems has thrown up some questions that need to be addressed very urgently for AI to be acceptable. What is acceptable AI, or what is expected from AI? One view is that it must have quality and values. Measures of quality depend on the paradigm that is employed for judging the utility of the system, and typically it is the market. Quality is local to the problem being solved and typically gets reflected in the utility function that the AI system attempts to optimize. An AI system can be deemed to have quality if it effectively solves the problem or performs the task for which it was designed.

An AI system can be made acceptable by adding values to it to ensure that it acts according to the accepted human values. This is typically referred to as the Value Alignment Problem (VAP). The problem of inculcating human values into an AI system is an overarching global concept. A typical solution may appear to solve a given problem and yet may not pass the value test. As long as the system is playing chess or Going, the rules of the game are the only values it needs to have or the only ethical considerations. But, in the case of a self-driving vehicle [6, 7], only the driving rules are not enough. Along with the driving rules, the AI system should be fed with values to make an acceptable autonomous AI system. In the absence of values, decisions taken by the AI system may not be acceptable. The fundamental issue is to decide what values to incorporate and how to incorporate them into an AI system. Can humans agree on a set of values that are universally accepted across all cultures? These are, of course, debatable. However, it is not difficult to arrive at a basic set of human values, a kind of a Common Minimum Program. Assimov's Laws [8] are a celebrated example of such a very basic Common Minimum Program.

Ethics or values can be defined in different ways, from religious to the philosophical point of view of what is meant by a good action [9]. The research area of machine ethics has recently emerged as a subfield of AI focusing on the task of ensuring the ethical behavior of artificial agents [10]. One of the issues is to define the exact set of values as people might have different cultures, come from different parts of the world, and have different socio-economic backgrounds.

The ethics of the AI system reflect the ethics of the designer/user. Ethics enable a final choice to be made among contending alternatives. It is not always possible to explain why a particular alternative has been chosen by the AI system. If any system is to gain significant control over the future, then it is imperative that the system be constructed to act according to the interests of not only its operators but all humanity and perhaps all sapient beings [11].

AI value alignment problem has two different aspects: normative aspect and technical aspect. The normative aspect of value alignment deals with the question of what values are to be incorporated into the AI system, whereas the technical aspect deals with the question of how to incorporate the values into the AI system.

CHAPTER 15

Cryptocurrency Portfolio Management Using Reinforcement Learning

Vatsal Khandor[1,*], Sanay Shah[1], Parth Kalkotwar[1], Saurav Tiwari[1] and **Sindhu Nair[1]**

[1] *Dwarkadas J. Sanghvi College of Engineering, Mumbai, India*

Abstract: Portfolio management is the science of choosing the best investment policies and strategies with the aim of getting maximum returns. Simply, it means managing the assets/stocks of a company, organization, or individual and taking into account the risks, and increasing the profit. This paper proposes portfolio management using a bot leveraging a reinforcement learning environment specifically for cryptocurrencies which are a hot topic in the current world of technology. The reinforcement Learning Environment gives the reward/penalty to the agent, which helps it train itself during the training process and make decisions based on the trial-and-error method. Dense and CNN networks are used for training the agent to taking the decision to either buy, hold or sell the coin. Various technical indicators, like MACD, SMA, *etc.*, are also included in the dataset while making the decisions. The bot is trained on 3-year hourly data of Bitcoin, and results demonstrate that the Dense and CNN network models show a good amount of profit against a starting balance of 1,000, indicating that reinforcement learning environments can be efficacious and can be incorporated into the trading environments.

Keywords: Actor-Critic, Bitcoin, Cryptocurrency, CNN, Portfolio Management, Reinforcement Learning.

INTRODUCTION

Cryptocurrency is an asset that is digital and is considered a medium of exchange. No physical form exists from cryptocurrency, like that of money, which is in the form of paper. They exist in the form of different types of coins. Some examples of coins are bitcoin, Ethereum, Cardano, Polkadot, *etc.* These coins are obtained by a process called mining.

* **Corresponding author Vatsal Khandor:** Dwarkadas J. Sanghvi College of Engineering, Mumbai, India; E-mail: khandorvatsal@gmail.com

Gyanendra Verma & Rajesh Doriya (Eds.)
All rights reserved-© 2023 Bentham Science Publishers

The ownership records of each coin individually are stored in a ledger that exists in the form of a computerized database. Since the transactions of these coins need to be secured, it is stored using strong cryptography. Also, strong cryptography is used to verify the transfer of coin ownership, and the creation of additional coins is controlled as well. Central authorities do not issue cryptocurrency. Cryptocurrencies typically use decentralized control for getting issued. When it is created, or rather in crypto terms, when it is minted prior to its issuance or before it is issued by a single issuer, that is, that coin belongs to that particular user; it is generally considered centralized. Cryptocurrencies work on blockchain technology, which is decentralized. The crypto transactions are managed and recorded by many computers onto which the blockchain technology is spread.

Hence, blockchain as a technology is preferred due to its security. Portfolio management is the control of an organization's projects and programs. Selection and prioritization are also made under portfolio management in line with the capacity to deliver for the company and its strategic objectives. The final goal is the maintenance of business as usual and to balance the implementation of change initiatives while optimizing ROI (return on investment). In simpler words, it involves a technique of analyzing and selecting a group of investments that meet the long-term financial objectives of a client, a company, or an institution. Also, risk tolerance is taken care of. Some of the fundamental objectives of portfolio management are capital appreciation, risk optimization, protecting earnings against market risks, ensuring flexibility of portfolio, allocating resources optimally, maximizing return on investment, and improving the overall proficiency of the portfolio, as demonstrated in [2]. Also, different types of portfolio management are active, passive, discretionary, and non-discretionary portfolio management. A portfolio can be managed in three different ways. The first is asset allocation, in which the investors put money into assets. These assets are both volatile and non-volatile. Money is put in such a way that, at minimum risk, we get substantial returns. The second one is diversification, wherein the portfolio of the investor is well-balanced. Also, the portfolio is diversified across different investment avenues. In doing so, the collection of the investors is revamped significantly by achieving a perfect blend of reward and risk. This, in turn, helps to generate risk-adjusted returns over time and also helps to cushion risks. The third method is rebalancing, in which investors rebalance the ratio of portfolio components to yield higher returns at minimal loss.

Reinforcement learning is a subset of machine learning where it is concerned with how intelligent agents will take actions in an environment so that the cumulative reward is maximized. It basically focuses on taking a suitable action that would, in turn, maximize the reward in a particular situation. Reinforcement learning is employed by various software and machines to find the best possible behavior and the best path which should be taken in a specific situation. Reinforcement learning is one of three basic machine learning types, the other two being supervised learning and unsupervised learning. The environment is typically stated in the form of a Markov decision process (MDP). The different types of reinforcement learning are associate, deep, safe, and inverse. Also, the different reinforcement learning algorithms as in [1] are Q-Learning, DQN (Deep Q Network), DDPG (Deep Deterministic Policy Gradient), A3C (Asynchronous Advantage Actor-Critic Algorithm), *etc.*

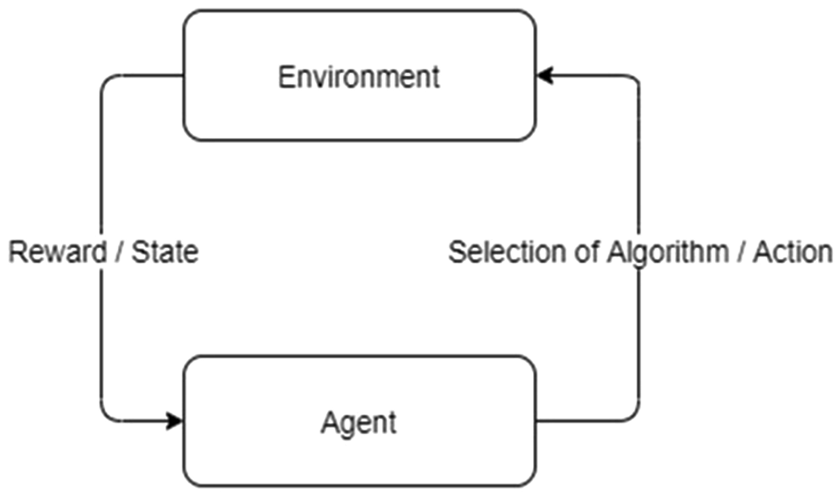


Fig. (1). Basic Block for Reinforcement Learning.

RELATED WORK

Various different machine learning techniques to perform algorithmic trading are observed. Author [1] proposes a novel trading approach that uses a Markov Decision Process (MDP), and to improve the performance of Q-learning, the author augments MDP with an estimate of current trend information and sentimental analysis of news articles. The implemented system performs sentimental analysis using Word2Vec and finds the trend analysis using Neural Networks, and they formulate the statement as an MDP and solve this MDP using Reinforcement learning. Though, the evidence on the use of a live dataset is still widely debated since the author uses only a 5-year historical price dataset. And for the sentimental analysis model, the author uses 10-years news articles obtained from Reuters. The author concludes the result based on the term “Sharpe ratio.” The author also captures the current trend information and formulates the augmenting trend information into an MDP. The author also uses 6 different technical indicators: SMA, Moving Average, Stochastic K, Stochastic D, RSI, and

SUBJECT INDEX

A

B

C

D

E

All rights reserved-© 2023 Bentham Science Publishers

N

O

P

R

S

T

U

V

W

X

Gyanendra Verma

Gyanendra Verma is currently an Assistant Professor (senior grade) at the Department of Information Technology, National Institute of Technology Raipur, India. He completed his B. Tech. in 2006 from Harcourt Batlar Technical University (formerly HBTI) Kanpur, India and M. Tech. & Ph.D. from the Indian Institute of Information Technology (IIITA) Allahabad, India, in 2009 and 2016, respectively. His all professional degrees (B.Tech, M.Tech & Ph.D.) are in Information Technology. He has teaching and research experience of more than 12 years in the area of Computer Science and Information Technology, with a special interest in Image Processing, Speech and Language Processing and Human-Computer Interaction. He has collaborated actively with researchers in several other disciplines of computer science, particularly on Medical Imaging problems at the SILP laboratory, IIIT Allahabad. He has served on roughly fifty conference and workshop program committees and served as the Organizing Chair for MIND 2019. His research on the application of Wavelet Transform in Medical Imaging and Computer Vision problems have been cited extensively. He is a senior member of Institute of Electrical and Electronics Engineers (IEEE) and Association for Computing Machinery (ACM); reputed International Technical Societies.

Rajesh Doriya

Rajesh Doriya received his BE in IT from MITS Gwalior, India. He has done his MTech in IT with a specialization in Robotics from IIIT Allahabad, India. He has also received his PhD from IIIT Allahabad, India. He has got teaching experience of more than 12 years and presently he is working as Assistant Professor in the Department of IT, at NIT Raipur. He has also been associated with research projects at Institute for System Studies & Analyses (ISSA) lab of DRDO New Delhi, India, and the Division of Remote Handling & Robotics (DRHR) lab of BARC Mumbai, India.

He has authored more than 70 research articles including Peer Reviewed Journals, Conferences, and Book Chapters. He has organized International Conference MIND-2019 and MIND-2020 as Organizing Chair and Publication Chair, respectively. He has also organized a number of Short-Term Training Programs & Workshops in the fields of Soft Computing Techniques, Robotics, and Cloud Computing. He has delivered more than 20 invited guest lectures in/at many conferences/workshops/institutes. He has also served as a Member of the Board of Studies of Sharda University Greater Noida, ITM University Gwalior, and ITM University Naya Raipur, India. He is also a member of the research project review committee for the Chhattisgarh Council of Science & Technology (CCOST), Chhattisgarh, India. His research interests include Distributed Computing, Cloud Computing, Artificial Intelligence, Robotics, Soft Computing Techniques, Network Security, etc.