PROCEEDINGS: ALLIANCE SYMPOSIUM ON ARTIFICIAL INTELLIGENCE FOR MEDICAL SCIENCE - 2022

Friday, March 25,2022

Editor:
Dr. Shobana Padmanabhan

Center for Social Impact and Research
Alliance University Bengaluru

Proceedings

ALLIANCE SYMPOSIUM ON ARTIFICIAL INTELLIGENCE FOR MEDICAL SCIENCE

Friday, March 25,2022

Editor:
Dr. Shobana Padmanabhan

Center for Social Impact and Research
Alliance University, Bengaluru

Title:	**Proceedings: Alliance Symposium on Artificial Intelligence for Medical Science 2022**
Date of Conferene:	25 March, 2022
Editor's Name:	Dr. Shobana Padmanabhan
Published by:	Center for Social Impact and Research (CSIR) Alliance University, Chikkahagade Cross, Chandapura – Anekal Main Road, Bengaluru, Karnataka 562106.
Publisher's Address:	Alliance University, Chikkahagade Cross, Chandapura – Anekal Main Road, Bengaluru, Karnataka 562106
Printer's Details:	Eagle Prints and Graphics, Bengaluru
ISBN:	978-81-955963-2-4
First Edition:	2023 February

PREFACE

The one-day symposium on Artificial Intelligence (AI) for Medical Science aims to bring together researchers, practitioners and professionals from academia, industry, hospitals, clinics, and service centers to participate in thoughtful and informative presentations and discussions. On this platform, experts can discuss recent advances in the diagnosis, interventions and the impact of cutting-edge developments in the research and evidence-based practices.

Topics include but are not limited to the following:

1. AI-based Image Analysis
2. AI-based Audio Analysis
3. AI-based Video Analysis
4. Conversational AI
5. Markers for AI
6. Nutrition-based Diet and AI
7. Math, Optimization, Algorithms behind AI
8. AI for Autism and Spectrum Disorders
9. AI for Mapping of Learning Activities
10. Open problems in Medical Science
11. Research challenges in AI for Medical Science

Editorial Committee
Symposium on Artificial Intelligence for Medical Science

ORGANIZING COMMITTEE

Patrons

Dr. Premanand Shetty
Chancellor, Alliance University

Mr. Abhay Chebbi
Pro Chancellor, Alliance University

Dr. Anubha Singh
Vice-Chancellor, Alliance University

Dr. Samir Ranjan
Associate Pro Vice-Chancellor, Alliance University

Co-patron

Dr. Murthy Chavali
Dean-Research

Convenor

Dr. Shobana Padmanabhan,
Head - Center for Social Impact and Research,
Associate Professor - Department of Information Technology

Co-convenor

Dr. Vipin Prasad
Professor - Department of Science

Coordinators

Dr. Chetan J. Shelke,
Associate Professor - Department of Information Technology

Dr. George Easaw,
Professor - Department of Operations Management

Dr. Senbagavalli M,
Associate Professor - Department of Information Technology

Ms. Jahanara Shaik

Ph.D. Scholar

MESSAGE FROM DEAN (RESEARCH)

Dear Colleagues,

On behalf of the organizing committee, I warmly welcome you all to the One-day symposium on **"Artificial Intelligence (AI) for Medical Science (AIMS-2022)"** organized by the Centre for Social Impact and Research (CSIR), Alliance University, Bengaluru on the 25th March 2022.

Artificial intelligence (AI) is an interdisciplinary science, part of computer science involved with creating smart machines typically required beyond human intelligence capable of performing unprecedented tasks. AI and associated technologies are increasingly utilized in several areas and are beginning to be applied to medical, healthcare, pharmacy etc. Implementing such technologies, especially in the medical field can save time, reduce the administrative processes, potentially can transform diagnosis and treatment processes, drug development, medical documentation, customized/personalized medicine, interpret complicated medical and healthcare data with the utilization of machine learning algorithm, remotely diagnosing patients, and overall patient monitoring and care.

Nowadays, Artificially Intelligent systems are used extensively in the medical sciences/healthcare sector. The application of artificially intelligent systems in healthcare for use by the general public is relatively unexplored. AI looks well-positioned to revolutionize medical science and its future in healthcare seems very bright and promising.

The Convenor & Head, Centre for Social Impact and Research and Co-convenor & Head, Department of Science have

taken a great theme to conduct this one-day event suitable to the situation. The event has already taken shape in a short time to be the best platform to showcase the student talent, networking opportunities, and collaborations with lead speakers. I add my best wishes for a successful and fruitful symposium. I look forward to learning the latest applications of AI with experts from industry, government, and academic scientists on these topics. Looking forward to an excellent symposium with experts from different countries sharing their new and exciting results.

With best regards,

Prof. Dr. techn. Murthy SSS CHAVALI Yadav

TABLE OF CONTENTS

Future of Artificial Intelligence (AI) Technologies in Medical Sciences

Keynote Speech

Chitharanjan Billa

Software Architect | Data Scientist | Entrepreneur |
Lead – Centre of Excellence
22nd Century Technologies Inc., USA

Health care is facing tremendous challenges, and with the arrival of COVID in 2020, the challenges only multiplied. With the exponential growth of patients, limited infrastructure, and dwindling number of health care support staff, it's becoming increasingly difficult to address these needs both for the government agencies and patient communities. Using Artificial Intelligence (AI) in medical sciences and other patient related services would be an efficient helper function to alleviate the burdens mentioned above.

In today's symposium, an effort would be made to unpack the existing AI usage in medical sciences, challenges this technology is facing and discuss the future of AI in medical sciences. Even before the arrival of COVID, many

health organizations, pharmacies, and other medical related industries are leveraging machine and deep learning to a large extent such as new drug discoveries, prevention, and preparedness strategy to mitigate pandemic and other communicable diseases.

AI technology is also lowering costs for the patients by reducing the hospital visits and effectively helping the administration staff in scheduling. It is estimated that more than 70% of the health care institutions are expected to utilize AI/ML/DL services in the coming 5 years.

One common scenario we see in the medical field is the erroneous recording of patient's critical data. This is not only a huge burden on patient to deal with inaccurate diagnosis, but it also consumes a lot of administrative time. AI technology, if used for automation, can rectify this issue thus improving favorable outcomes for patients.

AI goes beyond attending administrative tasks; this technology is widely used as tool for diagnoses. While the theory of AI replacing a human physician could be considered impossible, AI can make accurate and faster diagnoses and produce slightly more accurate results than a radiologist as per studies. Another promising field is the use of AI in genomics, not only to diagnose, but also to develop new drugs and vaccinations.

Currently, there are several challenges faced by the implementing institutions when trying to adopt AI. The biggest challenge is the significant budget allocation for implementing AI. Most of the budget allocation for AI implementation goes to hire talented people, and for their talent to produce Machine-Learning/Deep-Learning algorithms. Talented people who have a thorough understanding of Math branches such as Algebra, Calculus, Statistics and Probability are rare to find. There should be

particular emphasis on training new data scientists with appropriate math skills.

In addition to the above challenges, there are other technical and logical issues such as detecting outliers and the decisions to do when outliers were encountered. If an AI algorithm that is trained to identify cars and trucks comes across a cat, then it can typically detect that the cat — which is an "out-of-distribution" example — as an outlier. An emphasis should be put on building AI systems that can safely make decisions when encountered with such "out-of-distribution" anomalies. Safety should be the utmost priority when designing AI algorithms. Ethical issues while using AI should be a given priority. Future of AI technology in medical sciences looks very promising although few challenges exist, which could be mitigated over time.

Autistic Spectrum Disorders-Diversity and Prospects for AI

Invited talk

Dr. B. Anand

Professor and HOD, Department of Psychiatry
Malla Reddy Institute of Medical Sciences Hyderabad

Introduction: Autism Spectrum Disorder (ASD) is characterized by major Impairment in social functioning, emotional, and communication skills, Behavioral abnormalities, and sensory integration. Most autistic people can't tell, just by looking at your face, whether you are happy or sad, cross, or friendly, relaxed or worried.1 in 44 children has an autism spectrum disorder (CDC, 2021 Dec). It is more common in boys than girls (4:1), ASD occurs in all racial, ethnic, and social groups. In India it is1%. To diagnose ASD the development of a child needs to be assessed from information provided by parents and observations. In India, Developmental Monitoring at 18 Months–Modified Checklist for Autism in Toddlers (M-CHAT) is commonly used and for severity Indian Scale

for Assessment of Autism (ISAA) with age range from 2 to 29 years.

Diversity of ASD: The diversity of ASD in Clinical manifestations, Intelligence levels, differential diagnosis, wide variety of Brain area changes as evident from Brain Imaging. Default Mode Network (DMN) in brain have been seen, with local hyperconnectivity and general hypoconnectivity at distances. There are at least 85 different genes involved in ASD. Glutamatergic Synaptic dysfunction function and encoded proteins defects are seen in Autism. Intrauterine environmental factors are gaining increasing importance There are about 400 different treatments for ASD. There is no known cure and prognosis with good outcomes only is in 10–15% of cases.

Application of AI in ASD: A dense feedforward network provides the best results among the 12 models with an AUC value of 80%. Higher resolution of images (7 tesla) will let the images with texture analysis be more informative for predicting ASD[1]. In an EEG 3 to 36 months follow up study[2], Support vector machine (SVM) method with radial basis functions and Nonlinear analysis of EEG signals, significant differences were noticed in children who develop ASD, as early as 3 months of age. Developmental trajectories of ASD group diverged from the Low-Risk Controls group at early ages in the left temporal and right temporal-parietal regions, and diverged later, after 18 months, in frontal regions. Severity scores of ASD were significantly correlated with actual scores using EEG measurements taken at 3 months of age.

Prospects for AI in ASD: Behavioral Measures, Acoustic performance, Eye Tracking and Other Sensors, EEG and Electrophysiology, Gene Data, Neuroimaging and Treatment Assistance.[3]

Challenges in AI and ASD Interface

1. ML algorithms and the dimensionality reduction techniques are data-centric; they are independent of the conceptual basis of ASD assessment instruments.
2. The real-life implementation of AI and ASD assessment systems needs to be realized.
3. Public dataset needs to be continually expanded to avoid inappropriate studies due to insufficient data.
4. Researchers should have a detailed knowledge of the MRI, EEG, Eye Tracking etc. to eliminate errors.
5. Longitudinal data and Time series data from infancy to onset of ASD would yield better results.

Data Repositories with Links

- UCI https://archive.ics.uci.edu/ml/datasets.php
- ABIDEI http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html
- ABIDEIIhttp://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html
- NDAR https://ndar.nih.gov
- AGRE https://www.autismspeaks.org/agre
- NRGR https://www.nimhgenetics.org
- GEO https://www.ncbi.nlm.nih.gov/geo
- SSC https://www.sfari.org/resource/simons-simplex-collection
- Simons VIP https://www.sfari.org/funded-project/simons-variation-in-individualsproject-simons-vip
- Data repositories also used in Studies
- Boston Autism Consortium (AC)

- ASDTest: Kaggle and UCI ML repository
- Association of Parents and Friends for the Support and Defense of the rights of people with Autism (APADA)
- PASS app, Ondokuz Mayis University Samsun , and ASD outpatient clinics in Germany

References

[1] Deep learning applications for the classification of psychiatric disorders using neuroimaging data: Systematic review and meta-analysis. NeuroImage: Clinical 30 (2021) 102584

[2] S William J. Bos, Nature/ SCiEntiFiC REPorTS | (2018) 8:6828 | DOI:10.1038/s41598-018-24318-x

[3] Diagnostics 2021, 11, 2032

Harnessing technology to identify autism

Invited talk

Dr. Indu Dubey

Assistant Professor, Psychology
Academic Unit of Mental Health & Clinical Neurosciences, School of Medicine, University of Nottingham, Institute of Mental Health, Nottingham, United Kingdom.

Early diagnosis is key to early intervention in Autism Spectrum Conditions (ASC). Early interventions are associated with the best outcomes, and hence need to be prioritized. In low-resource settings such as India, lack of early detection of children at risk of ASC poses a significant barrier. According to latest WHO data on India, specialist (psychiatrist) to patient ratio can be as low as 0.30 per 100,000 people thus making the opportunities for early identification of the developmental disorders sparse. Hence, to fill the gap between the need for early identification and limited availability of the experts, it is important to develop alternative methods and tools that can be used by low skilled work force, are inexpensive, and are portable so that

they can be taken to the field without requiring any addition infrastructure.

In this talk, I shared the information about various autism screening and diagnostic tools available in India and rest of the world. I emphasized on discussing the technological adaptation of these tools as apps. I discussed the latest technology of using machine learning to identify gaze position in a video feed to evaluate social preference (or lack of it) in children with autism. I also discussed the use of technology to evaluate fine motor abilities by presenting dynamic stimuli to children on a tablet screen and recording their finger movements. In the end I emphasized the string need for similar technological advances to facilitate identification of red flag sign in autism using machine learning algorithms to evaluate free play interaction videos between children and parents or the speech production (atypicalities in speech) of autistic children.

AI in Healthcare

Dr. Chetan J. Shelke

Associate Professor, Department of Information Technology
Alliance University, Bangalore, India

AI has a great impact on health care industry. There has been a sudden growth of AI in health care, as a lot of data is available. That has made it easier to use AI in health care industry. Machine learning has capabilities to handle vast and critical data--process and analyze the data. However, for handling complex data, in addition to machine learning, we can use AI and deep learning, by implementing various algorithms on medical data. A major application of AI is in analyzing MRI scans. A large dataset of normal MRI scans takes a lot of time to analyze but deep learning and AI help analyze them faster and predict diseases faster. There are many AI devices available for giving warning for health conditions such as Apple's smart watch that monitors the human health by collecting and analyzing the medical data—the collected data is processed and analyzed by using machine learning and deep learning algorithms. Machine learning is subset of AI which enable machines to learn automatically. Deep learning is useful for solving complex problems and it uses the concept of neural networks.

COVID-19 Disease Detection Based on CNN Algorithm

Karan Ranjit Jagdale[1]
Chetan J. Shelke[2]
Rathnakar Achary[3]

Alliance College of Engineering and Design
Alliance University

[1]rkaranbtech19@ced.alliance.edu.in,
[2]chetan.shelke@alliance.edu.in
[3]rathnakar.achary@alliance.edu.in

Abstract — As the COVID-19 infection was spread to every corner of the community, the present research related to a method and system for its detection using CT scan images based on convolutional neural network. The system comprises of a database segment, for collecting a plurality of CT images; a processing segment, for feature selection and division of a dataset; a training and learning module, for training and testing the dataset. The method comprises steps of initially collecting the plurality of CT images. In the second step we have used Convolutional Neural Network (CNN) to excerpt attributes from CT scans.

Then performing feature selection on those images by using ConvNet or convolution neural networks. After the feature selection the plurality of images are divided into trainingg, testing and validation data which is used by the training and learning module. Finally, classification can be done by using classifiers (MLP and XGBOOST). This proposed system is effective in identification of COVID-19 infections in the second and future waves. Since the mutated virus in the second wave is not easily detectable by using the traditional method such as antigen test and RT-PCR.

Keywords**:** Convolutional Neural Network, Multi-Layer Perceptron(MLP), Computed Tomography(CT), K-Means Clustering, Optimization, Feature Extraction, Feature Selection.

I. INTRODUCTION

Coronavirus is a contagious respirational infection caused by acute coronavirus 2 (SARS-CoV-2). As a result of this there is a large death toll in the world, and the number of people affected is growing every day. It affected to the community and the economic conditions of the country are limitless. The first level of mitigation technique to prevent from the infection is social isolation, isolation, use of face masks. The treatment of this disease involves the rapid recognition of infected people through extensive testing and segregation of infected people. The ideal test mode would be a real-time reverse transcription polymerase chain reaction (rRT-PCR). The other tests such as Chest X-Rays and CT images also recommended by the doctors. However, in most instances, X-Ray and CT imaging can be helpful to radiologists in the initial diagnosis procedure [1].

New Coronavirus infections were first reported in Wuhan, China, and have since become more widespread since January 2020 worldwide. On January 30, 2020 the

World Health Organization (WHO) has confirmed that the occurrence of pandemic has become a worldwide public health emergency. One of the very common symptoms among the infected are high fever, respiratory tract congestion, mild difficulty in breathing, and cough. In addition, unexplained symptoms are fatigue, headaches, and shortness of breath (3%). The infected people usually start developing symptoms in a weeks' time after the infection. Animals can also transmit the disease and are also affected. Two major viruses were reported, namely Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS Coronavirus). Generally, the infected will suffer a serous respiratory problem and are naturally zoonotic.

In today's date, it usually takes up-to minimum of 24 hours to identify COVID-19 infection from the means of Qualitative RT-PCR, which by far has proven to be the most efficient way of tracking the infections, because the results obtained out this test is considered as quite accurate for the detection of SARS-CoV, but by no means is a fast method to track the spread of infections.

This makes it a very important aspect to detect the disease early and to place the infected person immediately in solitary confinement as there are no specific medications or specific emergency protocols available for COVID-19. RT-PCR faces high levels of non-fiction and is time-consuming. Low sensitivity testing of RT-PCR is unsatisfactory in today's epidemic situation. In some cases, those who may have been infected are not aware of the time and are not receiving appropriate treatment.

Compared to RT-PCR, Thorax Computer Tomography (CT) is likely to be a comparatively consistent, accurate, and fast-paced technology for the planning and testing of COVID-19, especially in those areas where COVID-19

has reached an epidemic stage. Most of the hospitals have the access to the CT technology, hence Thorax CT images is used for the analysis and detection of infection in asymptomatic or individuals showing mild symptoms. However, the separation of COVID-19 based on Thorax CT need a radiology specialist, which is a very sluggish process. Therefore, the automatic analysis of Chest CT images is beneficial towards saving valuable time for medical professionals. This helps to overcome the delay in starting treatment. The remaining portion of this paper is organized as follows. A detailed literature study in section II; proposed research methodology in section III, and section IV explain the result analysis followed with conclusion and scope for future enhancement in section V.

II. LITERATURE REVIEW

Several studies are carried out to analyze the spreading of COVID-19. The paper [2] describes, corona virus outbreak is due to the spread of COVID-19, is an RNA virus. Social isolation and wearing surgical mask are only the first level of protection to safeguard from the virus infection. The research result published in [6] is to create an awareness about the disease and helps to aid further research. The works in [3] proposed a machine learning model based on batch normalization and Mish function to detect COVID-19 cases from CT scanning. In this proposed system, two datasets were employed to gauge the proposed model. The system achieved an accuracy of 99.03.

The result presented introduces a network with kernel-based machine learning applications. The authors presented an AI-based framework in [4], which includes a smartphone enabled sensor with self-reported questionnaires. The initial result obtained confirmed the feasibility and efficiency of the solutions proposed. The framework demonstrates an

accuracy of 79% during the preliminary diagnosis process using ANN model.

In [5] the authors analyzed six different aspects related to corona virus and addressed six research questions comprehensively. This includes the virus growth trends in most infected countries such as India, forecasting the spreading of virus in the next few days and the mechanisms to minimize the spreading in the community.

Researchers in [7] proposed deep CNN model to diagnose the corona virus infection with a classification accuracy of 95%. Some researchers proposed COIVD-19 detection methods for multiclass classification [7].

In [9] authors proposed deep learning algorithm for prediction. In paper [10] a combination of three binary decision tree with CNN architecture in implemented on the PyTorch framework. This enhanced the accuracy up to 98%.

III. PROPOSED METHODOLOGY

It is an object of the present disclosure, which provides a method and system for COVID-19 infection detection using CT Images based on deep learning. The divesting effect of corona virus is considered one of the most predominant health crises of this century. Many famous organizations across the world came together for this tragedy and developed various Deep Learning models to better identify COVID-19 in Thorax CT scans. One such example was, the University of Waterloo, along with Darwin AI - the department's first startup, generated a database called 'COVIDx' containing 13,975 images of 13,870 cases and developed the Deep Learning model 'COVID-Net'.

In this study, COGNEX's Deep Learning Software, VisionPro Deep Learning, was used to distinguish these

Chest X-rays from the COVIDx database. The results obtained out of this are analyzed and compared with the results of COVID-Net and other models. In an embodiment of the present invention is, "The covid-19 Disease Detection Using CT Scan Images based on Deep Learning". The system comprises of a database segment, for collecting a plurality of CT scan images; a processing segment, for feature selection and division of a dataset; a training and learning module, for training and testing the dataset.

Deep learning tools are called black boxes as we are not able to elucidate why a particular model segregates images into distinct categories. This issue is solved by testing VisionPro Deep Learning with two settings, first, by determining the whole image as a Region of Interest, and second, by separating the lungs in the first step, and then making the separation step only for those who are separated lungs, instead of using the entire image. VisionPro Deep Learning Outcomes: In the whole picture as ROI it reaches a total F score of 94.0%, and in divided lungs, it gets 95.3% F points, better than COVID-Net and other provinces open source art models of art.

In an embodiment of the present invention, the covid-19 disease detection using CT scans based on deep learning. The analysis of a CT scan is complicated and necessitates the assistance of an expert. This assignment has several problems, including viewpoint variation, scale variation, intra-class variance, image deformation, image occlusion, lighting conditions, and backdrop clutter etc.,

Proposed system

In an embodiment of the present disclosure, the method of corona virus detection using CT Images based on deep learning comprises of following steps: collecting plurality of CT scan images of a plurality of patients at a node,

performing feature selection, adopting feature selection to form sample data, training the neural network using the training data, the convolution neural network using a deep learning model after division through the training data, testing the trained convolution neural network by using the test data; implementing the trained convolution neural network for predicting result of COVID-19 after testing.

In an aspect of the present disclosure, the first step in the method disclosed is collecting, by the database module, a plurality of CT scan images of a plurality of patients at a node. Wherein a node is a computerized device, mobile or computer handset. The database module can be any kind of memory of the node.

Convolutional neural network (CNN), the neural network model is a special type of deep learning algorithm that is designed to take input as two-dimensional image data, although there may be also one-dimensional and three-dimensional data. The most often used neural network model for image categorization is convolutional neural networks (CNNs). The essential premise of CNNs is that a local knowledge of an image is sufficient. The key feature of CNN is that, the number of parameters are independent of the original image size.

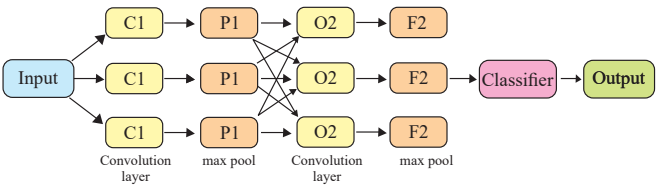


Figure 1: Basic Convolutional Neural

The study begins with the collection of primary datasets which further comprises of two classes. The first class belonged to chest CT image of the corona virus infection

with confirmed cases and the other class of images belonged to the normal people without the covid-19 disease.

Furthermore, in the next phase of the statistical data with the help of medical professionals, the dataset will be analyzed and CT scan which is not clear in terms of quality and diagnostics parameters will be removed accordingly. This process helps to generate clean data set. In the third phase, the dataset will be augmented by using standard augmentation techniques to increase its size. Furthermore, the results of the dataset of the third phase will be used to train the model of next phase.

Convolution CNN layer is the core building block, and it occurs most calculations. It requires several components, which are the input data, filters, and a feature. Let us assume that the input will be a color image, which is a three-dimensional matrix of pixels. This means, having an input three dimensions, height, width and depth, which corresponds to the RGB image. We also have a feature detector, referred to as a kernel or a filter, which in the moving image receptive field, check the presence of features. This process is called a convolution. Wherein the weight detector is a two-dimensional (2-d) array that represents part of an image.

While they can vary in size, the filter size is typically 3×3 matrix, which determines the size of the receptive field. The filter is then applied to a region of the image and calculates the dot product between the input pixel and filters. This dot product is fed to the output of the array. Then, the filter is moved by one step, the process is repeated until the entire image scanning core. From the product of a series of inputs and the filter is called a final output characteristic diagram, Fig activation, or a convolution function.

Further, the output value of each feature in the figures does not have a value of each pixel is connected to the input image. It only needs to be connected to the receptive field, where the filter is applied. Since the output of the array need not be directly mapped to each input value, the convolution (and pool) layer is typically referred to as "linking moiety" layer.

However, this feature may also be described as a local connection. Convolutional neural network is based on the results of research in neuroscience. They are the nodes of a layer of artificial neurons. These nodes are calculated and weighted input, and return an activation function of FIG. This is part of the convolution neural network. Each node in one layer is defined by its weight value. When you have some data to a layer, such as an image, it requires several pixel values and selected visual characteristics. Using data on CNN, return each layer activation map. These maps point out important features in the dataset.

If you give CNN the image, it will be noted that, like the color characteristics based on pixel values, and you activate the feature. While considering the image, CNN will first find the edges of the image. Then, this image will be passed to the next layer. This layer starts to detect things like the angle and color group. Then the image resolution will be passed to the next layer, so the cycle continues until you make predictions. Since the layer to be more explicit, this is to be called as maximum pool. It is only in the active layer of FIG Returns the most relevant features.

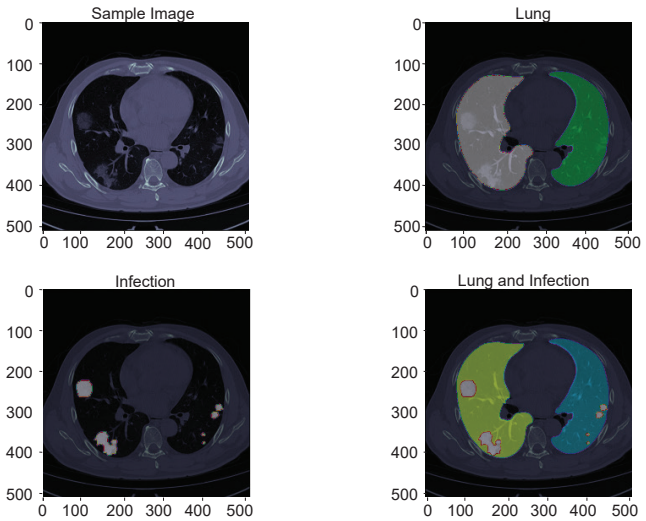


Figure 2: *Sample, Masked Lung Infection and Lung Infection Images*

This is to be delivered to each successive layer until you reach the last layer. Input, the random noise is Gaussian and values may be sampled and supplied from the distribution network to the power generator, and generates an image. The image produced is compared with the identification of the real image, and trying to figure out if a given image is positive or negative. Note that, in the feature detector right weight, as it moves through the image, which is also referred to as shared parameters remain fixed.

Some parameters, such as the weight value is adjusted during the training process by back propagation gradient descent in. However, there are three ultra-volume parameters of neural network training its effects needed to be set before the start of output. After each convolution, CNN applying linear rectifying means (RELU) converted to a characteristic graph of the nonlinear model is introduced.

As we mentioned earlier, the first layer may further convolution in accordance with convolution layer. When this happens, CNN structure of the pixel can be seen as within the receptive field layer delamination can be pre-layer shows a combination thereof representative of a portion of the higher-level pattern, characterized thereby hierarchy in the CNN. The final layer of a CNN is the predicted value is determined based on the classification layer on the activation map. If you pass a handwriting sample to CNN, the classification level will tell you what the letters in the image. This is a self-car use, to determine whether the object is another vehicle, a person, or some other disorder. CNN is like many other training machine learning algorithms.

IV. RESULT ANALYSIS

We use a U-NET architecture for segmenting the images. It is essentially a convolution neural network which has been refined for the purpose of segmenting medical images. The U-NET architecture has a greater amount of feature channels as compared to a normal convolution network.

The system includes a training section with training data to train the algorithm constraints and performance. The feature extraction of the model proposed is performed in the testing section.

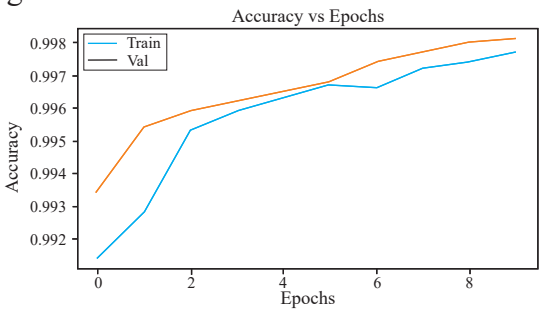


Figure 3*: Accuracy vs Epochs Graph*

As a result, the model has achieved an accuracy of 99.81% in 10 epochs. We have plotted the values for Epochs vs Accuracy. From these plotted values as in figure (3), we observe that the overall accuracy improved significantly for every Epoch staring from 99.41% right up-to 99.81%.

V. CONCLUSION

This study is related to COVID19 testing using CT scan image is demonstrated by using the COVIDx dataset. The dataset used, consisted over 20 CT Scan images along with segmented images of lungs. We used our trained model for predicting the infection mask from the original CT Scan image.

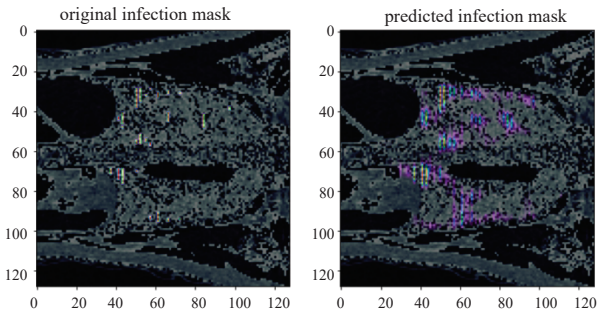


Figure 4*: Original Infection mask vs Predicted Infection Mask*

With this, we were able to establish a CNN refined for the detection of COVID-19 from CT-Scans with an accuracy of 99.81%. By any standards, this is not the exact solution, but is an effective method and has helped to lay foundation for providing positive results. As a future enhancement, the algorithm will be enhanced to analyse the severity of the COVID infection and to find more significant details from CT scan image to combat the pandemic.

REFERENCES

[1] Avishek Garain1 • Arpan Basu1 • Fabio Giampaolo2 • Juan D. Velasquez3 • Ram Sarkar1," Detection of COVID-19 from CT scan images: A spiking neural networkbased approach", springer, Neural Computing and Applications 16April2021.

[2] Albaraa A Milibari, "Current Situation of Coronavirus Disease: (COVID-19) Review Article" iMedPub Journals

[3] SeyyedMohammad JavadiMoghaddam a,* , Hossain Gholamalinejad b, "A novel deep learning based method for COVID-19 detection from CT image", Biomedical Signal Processing and Control 70 (2021) 102987

[4] Hayat Khaloufi 1 , Karim Abouelmehdi 1, Abderrahim Beni-Hssane 1, Furqan Rustam 2, Anca Delia Jurcut 3,* , Ernesto Lee 4 and Imran Ashraf 5," Deep Learning Based Early Detection Framework for Preliminary Diagnosis of COVID-19 via Onboard Smartphone Sensors" Sensors 2021, 21, 6853.

[5] Rajan Gupta1, Saibal K. Pal2 and Gaurav Pandey3," A Comprehensive Analysis of COVID-19 Outbreak situation in India", https://doi.org/10.1101/2020.04.08.20058347;April 11, 2020.

[6] Abel Brodeur David Gray Anik Islam Suraiya Jabeen Bhuiyan, "A Literature Review of the Economics of COVID-19" ZA DP No. 13411 Li et al. J Biomed Sci.

[7] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine, p. 103792, 2020.

[8] A.I. Khan J.L. Shah M.M. Bhat "Coronet: A deep neural network for detection and diagnosis of COVID-19

from chest x-ray images," Computer Methods and Programs in Biomedicine 2020 105581.

[9] M. Karim, T. Dohmen, ¨ D. Rebholz-Schuhmann, S. Decker, M. Cochez, and O. Beyan, "Deepcovidexplainer: Explainable covid-19 predictions based on chest x-ray images," arXiv preprint arXiv:2004.04582, 2020.

[10] S.H. Yoo, H. Geng, T.L. Chiu, S.K. Yu, D.C. Cho, J. Heo, M.S. Choi, I.H. Choi, C. Cung Van, N.V. Nhung, B.J. Min, H. Lee, Deep learning-based decision-tree classifier for COVID-19 .

[1] diagnosis from chest X-ray imaging, Frontiers in medicine 7 (2020), https://doi.org/10.3389/fmed.2020.00427.

[2] Rathnakar Achary, M Naik, T Pancholi, "Prediction of Congestive Heart Failure (CHF) ECG Data Using Machine Learning" Intelligent Data Communication Technologies and Internet of Things, https://link.springer.com/chapter/10.1007/978-981-15-9509-7_28.

CT Scan Error Diagnosis Using Deep Learning and Convolution Neural Network

Richard Martin

Student, M.Tech. in Data Science
Alliance University
mrichardmtech2020@ced.alliance.edu.in

Santosh Raghavendraraju

Student, M.Tech. in Data Science
Alliance University
rsantoshmtech2020@ced.alliance.edu.in

Dr Shobana Padmanabhan

Associate Professor
Dept. of Information Technology
Alliance University
shobana.padmanabhan@alliance.edu.in

Abstract — Fault diagnostics and prognostics are essential issues. Hospitals are under tremendous pressure to maintain unpredictable interruptions, system failures, and safety issues to a minimum which necessitates identifying

and eliminating potential issues as quickly as possible. Intelligent problem diagnosis is a promising technique because of its capacity to handle gathered signals quickly and effectively while also offering reliable diagnosis findings. Numerous authors have validated deep learning and machine learning approaches for identifying bearings failures, the findings have mostly been confined to tiny training and testing datasets, with the input data modified to achieve high accuracy. In this article, original data of an accelerometer sensor was loaded into a unique periodic sequencing prediction algorithm that led to an edge fault diagnosis technique. We utilize identical frequency patterns as inputs to an innovative deep neural Long-Short-Term-Memory Recurrent Neural Network to diagnosis bearings insufficiency at excellent accuracy in the least time period. Without the use of data adjustment, the technique would acquire the maximum level of competence in the industry. The fault diagnostic method's efficacy and applicability are demonstrated by comparing the findings to those of other intelligent fault detection systems using two widely known benchmark real vibration datasets.

Keywords— *Deep learning, Imps, Cwru, Predictive analytic*

I. Introduction

The use of machine learning or deep learning in numerous science disciplines have grown dramatically. Intelligent defect detection is one of the topics that has sparked a lot of attention and is being applied in real-world situations. A most difficult part of adopting machine learning methods to diagnose bearing failures is designing a system architecture that provide appropriate diagnostics reports in a short time. [2]. Signal processing technology is used to identify data-driven intelligent bearing flaws. These impulses are referred

to as "vibration signal", inertial sensors or bandwidth amplifiers are being used to analyze the "motor pulses" [3]. For its more precise results, the frequency range had attracted a lot of attention in the research [4]. To achieve the highest degree of precision while using machine learning and deep learning procedures in bearings condition monitoring, we must identify variables and afterwards use those toward the process of learning. Terms of low statistics, comparable methodologies typically produce adequate outcomes [8]. Yaqub et al., for reference, utilize KNN for bearings imperfection diagnosis and evaluation on such a shorter set of data, but also relatively high ability to gather (HOC) and wavelet (WT) in predetermined data analysis, despite their lack of accuracy. Hu et al. employ SVM in [10] in a similar way that the data information has indeed been pre-processed, as well as the records being minimal.

Huge amounts of data were accumulated in recent years of accelerated establishment of innovative monitoring systems, as well as the preponderance of a documented standard Machine learning algorithms have constraints in establishing evaluation methods on these data [11]. As a result, conventional approaches had given way to more complicated tasks such as deep learning models (DNN),etc. Eren et al. utilize a one-dimensional deep neural network to processed data in time series forecasting [12].

To improve efficiency, the supplied signal is analyzed, compressed, and standardized. In [11], Zhang et al. state how they can forecast time sequences with extreme accuracy employing DNN, however they do not provide the architecture of their recommended DNN network. An equivalent trouble comes in [2,] wherein Mao et al. profess to have obtained high precision via applying a customized classification algorithm. Since the designers mainly provide accuracy percentage (not testing accuracy) but do not

provide a reasonable infrastructure for its neural framework, subsequent reproducibility was constrained. We also found across all these edge detection, including [13] and [14] that examined short-term memory networking in bearings fault detection simultaneously. The framework, and the techniques that are used to obtain that is our proposed model.

The following issues may be found in all past attempts at bearing problem diagnostics: 1. Characteristics are chosen or tweaked. 2. The dataset's breadth is limited, and it is incapable of providing full data on a large scale. 3. The accuracy is acceptable, but not sufficient to depend on the result on a big scale. 4. The neural network architecture is very vaguely explained, and the method for obtaining the stated accuracy is uncertain.

We use a CNN-LSTM network in this study to predict the periodic sequencing of data provided inside the time - frequency domain, with maximum quality, in the shortest amount of time. They need not pre-process or enhance the raw information. As a function, this framework could be used in every domain to retrieve the fundamental signal characteristics of a proper system. These results show that the proposed approach outperforms region articles for both the training and test datasets (databases), in significantly less time frame. We will provide infrastructure and measures that achieve the highest effectiveness of our proposed fault detection approaches.

II. Methodology

CNNs were widely used to develop feed-forward ANNs which are used to resemble the mammal sensory cortex [17]. Time-series classification and forecasting were promising possibilities for one-dimensional CNNs. Consequently, this system was already used in cutting-edge technologies

including early diagnosis, predictive maintenance, and anomalies' rapid detection [18]. Since our input was built up of vibration analysis (time series), the outcome of a convolutional layer is expressed as follows [19] at location x of the i^{th} layer's j^{th} conv layer.

A Pooling layer is typically used after one or more CNN layers to achieve invariance by decreasing the feature map resolution [20]. Every convolution operation correlates with the convolutional layer. Max-pooling is the most frequently employed pooling approach. Here e is the component of the nth patchwork size.

Recurrent networks are fully convolutional structures with a unique gating mechanism which modulates memory cell accessibility [21]. Though gating can block the rest of the networks from updating the information of memory elements for a few clock cycles, recurrent networks could retain the information and propagate defects for much longer than conventional machine learning algorithms. Hochreiter et al [21] developed the LSTM that efficiently represents periodic sequence and the associated long-range interconnections, better than regular RNNs. Each machine learning model is formed by three stage: an input stage, a forget stage, and an output stage. Everything comprises a nonlinear activation function which generates integers between 0 and 1.

If a value is 0, no data is transmitted through the gateway; if a value is 1, all the data is transmitted through the gateway [22].

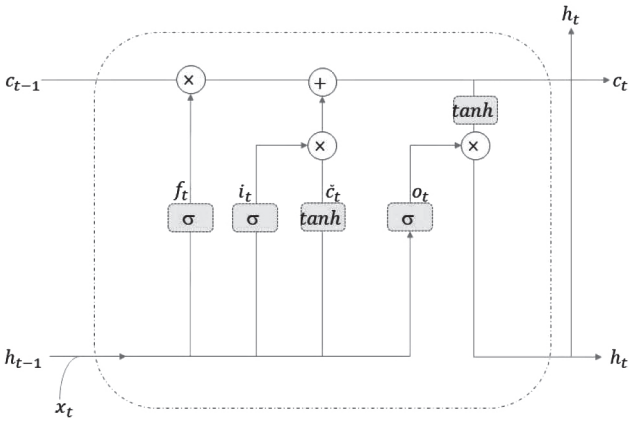


Figure 1. *The Long short - term memory component.*

CNN-LSTMs were utilized for voice commands and natural language in addition to certain other visible teaching processes [23]. Furthermore, CNNs and LSTM are effective algorithms for predicting temporal sequences [24]. While interacting with massive quantities of data or complicated spatial sequencing problems, the Deep network enhances predictions accuracy and precision [25].

Illustrate spatial sequenced assumptions, we analyze a model of the system within a temporal area described by "M" and "N" grids with "M" rows and "N" columns. Every grid-cell contains P readings that fluctuate across time. Consequently, a PxMxN convolution could be used to describe the amount of characteristics. As a result of separating the information into samples of comparable spatial duration, to obtain a sequence of convolution operation X1, X2,..., Xn. The temporally sequential prediction model's objective will be to improve the probability mean. [26]. Determine a most anticipated k^{th} sequencing of an observations based on the previous Jth Inspection.

Architectural & Instructional Techniques

In time-series, a strong 1-dimensional architecture occurs, including intimately correlated parameters (or pixels) that really are temporally close together. Local connections are always responsible while collecting and integrating local information before identifying spatial information [27]. LSTMs, on either side, might form relation between different elements [28], enabling devices can handle effectively. As a consequence, blending different network assists on research methodology. Convolutional networks have a de-noising functionality which reduces overall incidence of interference inside the learning experience and demands minimal pre-processing. Additionally, the Convolution layer is much more resistant against clustering that existing neural network models. An improved CNN-LSTM model is selected in bearings trouble diagnosis based on the preceding characteristics and multiple experiments conducted to build the best model that provides the greatest accuracy in the least possible time. It's important to mention that even if the CNNs and LSTMs were misplaced, and it will be contaminated by noise, but the information will be handled globally during first instance. The characteristics or properties of vibration analysis were obtained in the temporal domain.

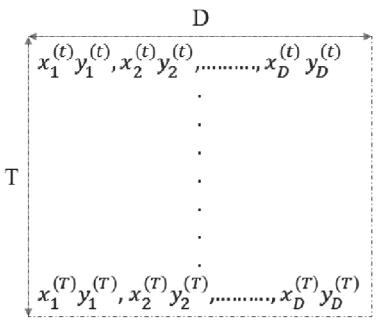


Figure 2. *Vibration signal statistics*

Determine the characteristics of a i^{th} orientations at time step t, wherein "D" describes the amount of bearing (test-cases) and "T" the testing duration.

Considering CRNN networks demands same size input transfer functions, first stage of segment the database sampling of width and length to provide to our CRNN model. The characteristics are classified into three **categories**: preparation, evaluation, and assessment. The model's hyper-parameters were determined by minimizing a cost function. The recommended architectural was made up of layers of LSTMs and single layer Part 3 investigates the planned infrastructure. To determine the optimal hyper-parameters towards effectively reducing the cost function. Parameters and LSTM with Twenty four layers, we were capable of achieving sufficient precision in our experiments. Every primary step is represented by a convolutional network, which efficiently eliminates fitting problem through reducing the connectivity across layers [30].

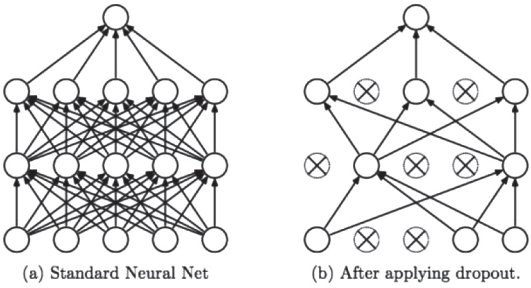


Figure 3. *Prototype on Dropouts Convolutional Network. A fully convolutional network featuring hidden layers can be seen on the left.*

Dropouts were added towards the connection on the left, resulted inside a limited connection [30].

3. Experiments

Throughout this section, we place our proposed technique to test using two benchmark functions: IMS and CWRU. The testing equipment designed to obtain vibration signals, failure categorization and Raw vibration signals used as inputs to our CRNN algorithms would be explained in the sections. Following that, the article identifies the datasets magnitude /time graphs, the aimed to develop our infrastructure, as well as the precision graphs and confuse matrix for every experiment.

4. IMPS Scan Dataset

A proposed model's effectiveness has been evaluated by using experimental datasets. illustrates the measurement device. As illustrated in data, there seems to be a spindle with multiple bearings. These machine were Rexnord ZA-2115 dual image sensor. Each bearing is connected in Cartesian coordinates to two high-precision accelerometers, allowing for vibration monitoring both in X and Y axes A conveyor system connects the driveshaft to an alternating - current (AC) motor that operates it. A longitudinal force of 2721.5 kg is supplied to a shafts and bearing by a pneumatic actuator. At a pace of 2000 revolutions per minute, the shaft revolves (RPM). A single file is produced for every 20480 data points at a sampling rate of 20 KHz (collected in one second). Whereas the CT scan were spinning, information is collected and recorded in records each 5 to 10 minutes. There really are 2156 files for every experiment. As a consequence, each test includes 44,154,880 data points in total. According to previous research on the IMS CT scan dataset initial test, there really are seven unique states of health during the assessment test [32].

- Unique
- Investigative

- Immediate failures
- Failure of the scans
- Stage 2 breakdown
- Tapered roller breakdown

Since the dynamic characteristics of certain regions are really comparable and can be characterized by signal analysis, to minimize computational effort as well as improve the performance with our learning algorithm, we decided to use these same classification with greatest significance for both fault detection and practical implementations.

Health Evaluation,

- Failure of the image scans
- Failure of a roller cameras

As suggested previously, the number of data points in first assessment is rather enormous, using this huge dataset as inputs in our learning approach has been both expensive and memory intensive. As a result, we picked 30 records randomly for every health class or phase. Following this, the data is concatenated, labelled, and prepared to be submitted to the deep learning model. These classifications were as continues to follow: 0-(Healthy), 1-(Suspected), 2-(Internal error), and 3-(Inner fault). Regardless of the fact that we'll be implementing a CRNN infrastructure, the input has to be a sequence of convolution operation that identical dimensions, as explained inside the previous subsection. With sampling frequency of 20kHz and a spindle speed 2000 RPM, 600 points per revolutions can be calculated (rotation period). For each sample, one-fourth of the rotation time, or 150 rows of data, is generated. Bearings vibration data both in A and B axes is included in each block of data. As a conclusion, each row contains the eight characteristics of

accelerometers of each bearings listed CA1 and CB1, CA2 and CB2, CA3 and CB3, CA4 and CB4.

Furthermore, every occurrence is a tensor of size (150x8x1), and the input tensor for every health state is a tensor of size (150x8x1) (4096x150x8). Table 1 showing the number of samples for each class. Each health status does have its own set unique resonance characteristics.

Stages	**No.of samples**	**Indication**
Healthy	4085	0
Suspected	4092	1
Internal error	4098	2
Inner fault	4097	3

Table 1 *– For each health status, both number of tests and class number.*

The data is classified into three categories: assessment, validation, and learning. To use a domain specific decisions are being made such that the number of observations inside the training and testing sets is integer and divisible by the batch-size. There seem to be a total of 16,384 samples throughout all four classes. As a conclusion, we designate 25% of the database to training combinations and 75% to evaluate settings. Determine the best architecture for our model, we evaluate various network with variable hyper-parameters.

The number of iterations are set to 50 for all performed tests, with the goal of achieving the highest training accuracy in the shortest time possible. Tensorflow, a Python library, is being used to create the simulation environment. The processor is an Intel(R) Core(TM) i7-8550U CPU with four cores and eight logical processors functioning at 1.80GHz and 1992MHz, correspondingly. RAM is 8GB.

Test quantity 12 provides the best results. Increasing the convolution layers, as predicted, results in a more comprehensive sample of the input [33]. Displaying the parameter tests 13 and 14, not only reduces test accuracy and also considerably increases computation time. In test 4, the ideal value for the LSTM synapses is identified. Based on the numerous attempts we performed, a most appropriate Table.2, demonstrates the system parameters in our developed framework. This incorporates a 24 layered Long short - term memory layer as well as a conv1D layer with 84 filters and a kernel size of 84. Differential evolution have the best keras optimizer functions.

1. CWRU

The test equipment consists of 2hp electric motor, a scanner a rings with an attached accelerometer, a torque transmitter, and a multimeter. In the experiment, SKF profound ball bearings 6205-2RS JEM were utilized. At 12 o'clock, the accelerometer was installed to the motor housing. At one frequency of 12 KHz, information for such drive-end-bearing testing were collected. The testing bearing were electro-discharged and create a specific spot imperfection with breakdown characteristics of 0.53mm at the inner raceway, rolling element, and surface roller cameras. The motor's frequency is estimated to just be 1700 rpm.

- Conventional
- Defective in the spherical
- Internal error 3 p.m.
- Surface error 3 p.m.
- Surface error at 6 p.m.
- Surface error at 12 p.m.

Methodology and techniques	**Pre-processing**	**Accuracy rate: IMS**	**Accuracy rate: CWRU**
KNN [9]	HOCs and WT	-	90.125%
SVM [10]	WP	62.5%	97.5%
SVM Combination [10]	WP	-	87.5%
SVM [34]	-	-	76.12-92.17%
DNN [35]	-	95.8%	93.1%
One Dimensional CNN[12] Enhancement of the processing unit	Centrally statistical linear resentation	95.2%	92.8%
Proposed Model	Sorting - Ruination - Validation	0.89	0.9789

Table 2. Vibration signals for Normal, Suspicious, Internal Error, and Spinning component fault IMS Bearings information

In order to implement an effective, the inputs to our CRNN model had to be a sequence of convolution layers of equal dimensions. With a sampling frequency of 12 KHz as well as a spindle speed of 1720 rpm, there are approximately 407 points every revolutions. We fixed the quantity of data points in each example at 202 in order to achieve the highest training and testing accuracy, which really is roughly twice a rotation.

In furthermore, every number of data with each general health is determined to just be divided by the number of observations each sample. Every example is a tensor, and the input tensor for every health conditions are same dimensions. Equivalent examples can be used to establish train, verification, and testing dataset. To utilize a domain specific the division and batches length in the LSTM model should be designed in such a way that the observations in the training and testing sets were integers and separated by the number of iterations. We allocate 50% of the dataset to learning and 50% to assessment. As just a result, for 50 iterations, the training and test accuracy rate were 1.0000 and 0.9977, correspondingly. The computation of 50 iterations requires 61 secs.

The transmission waveforms of a two classifications were virtually identical and the mistake can be disregarded since both errors are on the outside race. Considering that no data preparation, data selection, or manipulation was performed, the model's defect diagnostic strength for this test may be regarded as excellent.

3. DISCUSSION

Almost all previous techniques in the literature needed data pre-processing in some way Filtration, higher-order-cumulants, discrete wavelets, and wavelet packet transformation were a few of the applications. The high-dimensional returning characteristics are now used in combination with dimensionality reduction strategies, including principal components, to select the appropriate selection of characteristics. Since the dynamic characteristics of certain regions are really comparable and can be characterized by signal analysis, we decided to use these same classifications with the greatest significance in both fault detection and in realistic applications to reduce

the computation complexity and improve performance of our deep learning model.

To categories the selected characteristics, a variety of classifiers were used; however, we can see that preferences have developed over time extending from simple predictive learning techniques including such SVM and K - nearest neighbors to much more comprehensive learning methods including such Convolutional neural network and DNN. Almost all of these studies have one major flaw: they have exploited selective data or altered attributes to improve accuracy. In some cases, the altered properties may not accurately reflect the signal characteristics of the real system. As a result, this method limits the solutions' wide applicability.

Consequently, obtaining high-dimensional characteristics, and also any post-processing or feature extraction procedures which may be performed, could substantially increase the system's price and computational complexities. [12]. Despite the fact that several research have demonstrated good classification accuracy, their results are usually confined to tiny train and test datasets In comparison to the other publications listed in Table 6, we employed a larger data-frame of learning features and achieved greater accuracy without any data pre-processing or feature modification. Furthermore, several earlier research claimed to achieve high accuracy despite failing to disclose the architecture of their suggested network or a plausible path to accuracy [2,35]. The network design that was utilised to attain high accuracy was reported in this research, as well as the trials that went into selecting each individual piece of the proposed network.

4. CONCLUSION

This study comprehensively investigated the performance of a generic real-time CT Scan defect diagnostics. A CRNN classifier is used in the intelligent system, which is supplied by raw time-domain characteristics that are molded into tensors of time sequence. Using edge extraction of features approach, the raw addition to utilizing information is automatically and appropriately learned. In a short amount of time, the model could accurately identify the error while taking into consideration the enormous datasets. When compared to other ways, implementing the suggested method in a real context and on an industrial scale provides the following advantages. This technique is resistant to imbalanced datasets, this might be used to evaluate a greater and much more comprehensive set of data obtained. In comparison to previous publications in the literature, we were able to make a better accurate prediction in a shorter amount of time and with fewer epochs. No data preprocessing, predetermined modification, or controlled extracted features are performed because the framework is the final stage and can be generated directly from the raw vibration analysis. The computations become less expensive than certain techniques in the research, including database preprocessing and some complicated hierarchical structures.

Two standard vibration datasets are being used to test the CRNN classifier-based fault detection mechanism for bearing failure diagnosis. Experimental results validate the CRNN classifier's effectiveness and accessibility in defect diagnosis. Overall recognition accuracy rates for the IMS and CWRU bearing datasets were 97.13 percent and 99.77 percent, respectively. According the classification findings, the CRNN model acquired extremely discriminant information directly from raw sensor input information.

References

[1] Bonnett AH, Yung C (2008) Increased Efficiency Versus Increased Reliability. IEEE Industry Applications Magazine 14 (1):29-36. doi:10.1109/MIA.2007.909802

[2] Mao W, Feng W, Liang X (2019) A novel deep output kernel learning method for bearing fault structural diagnosis. Mechanical Systems and Signal Processing 117:293-318. doi:https://doi.org/10.1016/j.ymssp.2018.07.034

[3] Lessmeier C, Kimotho J, Zimmer D, Sextro W (2016) Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification.

[4] Bellini A, Immovilli F, Rubini R, Tassoni C Diagnosis of Bearing Faults of Induction Machines by Vibration or Current Signals: A Critical Comparison. In: 2008 IEEE Industry Applications Society Annual Meeting, 5-9 Oct. 2008 2008. pp 1-8. doi:10.1109/08IAS.2008.26

[5] Zhou W, Habetler TG, Harley RG (2008) Bearing Fault Detection Via Stator Current Noise Cancellation and Statistical Control. IEEE Transactions on Industrial Electronics 55 (12):4260-4269. doi:10.1109/TIE.2008.2005018

[6] Schoen RR, Habetler TG, Kamran F, Bartfield RG (1995) Motor bearing damage detection using stator current monitoring. IEEE Transactions on Industry

[7] Applications 31 (6):1274-1279. doi:10.1109/28.475697

[8] Eren L, Devaney MJ (2004) Bearing damage detection via wavelet packet decomposition of the stator current. IEEE Transactions on Instrumentation and Measurement 53 (2):431-436. doi:10.1109/TIM.2004.823323

[9] Samanta B, Nataraj C (2009) Use of particle swarm optimization for machinery fault detection. Engineering Applications of Artificial Intelligence22 (2):308-316. doi:https://doi.org/10.1016/j.engappai.2008.07.006

[10] F. Yaqub M, Gondal I, Kamruzzaman J (2012) Inchoate Fault Detection Framework: Adaptive Selection of Wavelet Nodes and Cumulant Orders, vol 61. doi:10.1109/TIM.2011.2172112

[11] Hu Q, He Z, Zhang Z, Zi Y (2007) Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble, vol 21. doi:10.1016/j.ymssp.2006.01.007

[12] Zhang R, Peng Z, wu L, Yao B, Guan Y (2017) Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence. Sensors 17:549. doi:10.3390/s17030549

[13] Eren L, Ince T, Kiranyaz S (2018) A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. Journal of Signal Processing Systems. doi:10.1007/s11265-018-13783

[14] Pan H, He X, Tang S, Meng F (2018) An Improved Bearing Fault Diagnosis Methodusing One-Dimensional CNN and LSTM. doi:10.5545/svjme.2017.5249

[15] Yoshimatsu O, Satou Y, Shibasaki K (2018) Rolling Bearing Diagnosis Based on Deep Learning Enhanced by Various Dataset Training.The Proceedings of the Symposium on Evaluation and Diagnosis 2018.17:109. doi:10.1299/jsmesed.2018.17.109

[16] J. Lee HQ, G. Yu, J. Lin, and Rexnord Technical Services (2017) IMS, University of Cincinnati. “Bearing Data Set”, NASA Ames Prognostics Data

[17] Repository. Center for Intelligent Maintenance Systems (IMS), University of Cincinnati

[18] Case Western Reserve University Bearing Data Center Website (http://csegroupscaseedu/bearingdatacenter/home)

[19] Alex K, Sutskever I, Hinton GE (2012) ImageNet Classification with Deep Convolutional Neural Networks.1097--1105

[20] Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman D (2019) 1D Convolutional Neural Networks and Applications: A Survey.

[21] Chen Y, Jiang H, Li C, Jia X, Ghamisi P (2016) Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks, vol 54. doi:10.1109/TGRS.2016.2584107

[22] Zhen Zuo BS, Wang Gang, Xiao Liu, Xingxing Wang, Bing Wang, Yushi Chen (2016) Learning Contextual Dependencies with Convolutional Hierarchical Recurrent Neural Networks. IEEE Transactions on Image Processing 10.1109/TIP.2016.2548241

[23] Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. Neural Computation 9 (8):1735-1780. doi:10.1162/neco.1997.9.8.1735

[24] Zhang Y, Hao X, Liu Y (2019) Simplifying Long Short-Term Memory for Fast Training and Time Series Prediction. Journal of Physics: Conference Series 1213:042039. doi:10.1088/1742-6596/1213/4/042039

[25] Bilgera C, Yamamoto A, Sawano M, Matsukura H, Ishida H (2018) Application of Convolutional Long Short-Term Memory Neural Networks to Signals Collected from a Sensor Network for Autonomous Gas Source Localization in Outdoor Environments. Sensors 18 (12):4484

[26] Yao H, Wu F, ke J, Tang X, Jia Y, Lu S, Gong P, Ye J (2018) Deep Multi-View Spatial-Temporal Network for Taxi Demand Prediction.

[27] Huang C-J, Kuo P-H (2018) A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities, vol 18. doi:10.3390/s18072220 26. Shi X, Chen Z, Wang H, Yeung D-Y, Wong WK, Woo W-c (2015) Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting.

[28] Lecun Y, Bengio Y (1995) Convolutional Networks for Images, Speech, and Time-Series.

[29] Zhang X, Chen F, Huang R (2018) A Combination of RNN and CNN for Attention-based Relation Classification. Procedia Computer Science 131:911917. doi:https://doi.org/10.1016/j.procs.2018.04.221

[30] Sainath TN, Vinyals O, Senior A, Sak H Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 19-24 April 2015 2015. pp 4580-4584. doi:10.1109/ICASSP.2015.7178838

[31] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting, vol 15.

[32] Ioffe S, Szegedy C (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.

[33] Qiu H, Lee J, Lin J, Yu G (2006) Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics.

[34] Journal of Sound and Vibration 289 (4):1066-1090. doi:https://doi.org/10.1016/j.jsv.2005.03.007

[35] Claessens BJ, Vrancx P, Ruelens F (2018) Convolutional Neural Networks for Automatic State-Time Feature Extraction in Reinforcement Learning Applied to Residential Load Control. IEEE Transactions on Smart Grid 9 (4):3259-3269. doi:10.1109/TSG.2016.2629450

[36] Wang Y, Liu F, Zhu A (2019) Bearing Fault Diagnosis Based on a Hybrid Classifier Ensemble Approach and the Improved Dempster-Shafer Theory. Sensors (Basel) 19 (9). doi:10.3390/s19092097

[37] Zhang R, Peng Z, wu L, Yao B, Guan Y (2017) Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence, vol 17. doi:10.3390/s17030549

[38] Zhou F, Yang S, Fujita H, Chen D, Wen C (2020) Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. Knowledge-Based Systems 187:104837. doi:https://doi.org/10.1016/j.knosys.2019.07.008

Depression Detection based on NLP & ML techniques using Text and Speech Recognition

Shanti Konda

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University Bangalore, INDIA
kshantiMTECH2020@ced.alliance.edu.in

Muralidhar Billa

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University Bangalore, INDIA
mbillaMTECH2020@ced.alliance.edu.in

Dr. Shobana Padmanabhan

Associate Professor - Dept. of Information Technology,
Alliance College of Engineering and Design
Alliance University Bangalore, INDIA
shobana.padmanabhan@alliance.edu.in

***Abstract*—**Suicide can be said as an act of taking one's own life voluntarily and intentionally. Suicide can be

prevented by identifying the condition of the person whether there are any suicidal tendencies and then by reporting it to the dearer or next to the kin. Detection of suicide ideation/ tendencies or depression of a person is done by using their speech and text from different languages with the help of Natural language Processing and Machine Learning techniques. Reporting is done by using an alert system in the form of email, call, SMS, and speech that is used to identify suicidal tendencies which can be of any language, and it is translated using Google Translator to English. In email and SMS, geolocation is sent as an attachment.

Keywords— *Suicidal Tendencies, Detection, Alert System, Speech Recognition, Translation*

I. Introduction

Death itself is unpredictable. Type of deaths can be categorized into accidental, homicidal, suicidal, and natural. Accidental deaths can be reduced by taking precautions, natural deaths are those due to health reasons or diseases like cancer, heart disease, etc.., which can be cured by treating and the invention of medicines. Reasons for suicidal deaths might be Depression, Traumatic Stress, Mental illness, Chronic pain and illness, Loss or Fear of Loss, Hopelessness, Social Isolation [10]. Attempting suicide due to the above-mentioned reasons can be prevented with the help of a psychiatrist or a psychologist. But this can be done only when it is identified beforehand.

In the 20th century, humans are getting attached to technology in every aspect. Smartphones are one of the devices that humans get addicted to, irrespective of age which leads to giving importance to technology rather than humankind. This usage of smartphones helps identify suicidal tendencies of a person through Artificial Intelligence.

In NLP techniques, Sentiment Analysis is one of the best approaches to predict the sentiment in a sentence provided. For example, Twitter data is widely used to determine the sentiment of a given statement. Based on this technique, thorough checking is done for suicide or depression datasets, and during Exploratory Data Analysis, it is observed that 50 percent of the deaths that occurred in the past 2 decades are suicides through Reddit data. These are the reported deaths that are known as per the data, and many of them are unknown. Every year more than 700 000 people die due to suicide. A prior "suicide attempt can be said as the leading risk factor for suicide in the general population". The 4th prominent cause of death in 15 to 19-year-olds is suicide. "77% of global suicides are observed in low-income countries and middle-income countries rather than high-income countries. Pesticide ingestion, hanging, and firearms are commonly used methods for suicide globally"[1].

Prevention of suicide remains a challenge, especially among teenagers and veterans. According to the "United States Department of Veterans Affairs (VA)", on average, in a day, around 17 veterans are dying from suicide, and that rates increase gradually [2].

"Depression detection is a well-known risk factor for suicide" and that can be said as a predictor of long-term or immediate suicide attempts and deaths [3,4]. In a human lifespan, ending life is not a solution since there are many ways to overcome the pain and suffering like sharing them with a trustworthy person. But the problem here is, the persons having depression might not share their burden, and to solve that, this idea can be used. The idea to identify depression is from text/typing through the keyboard and voice/speech through the microphone.

For identifying depression or suicidal ideation, speech features are considered which can be classified into acoustic features and linguistic features in which Google Cloud Speech API is used to convert speech to text and audio

files which are extracted using pyAudioAnalyisis python library are used for audio analysis [5] but audio files and acoustic features are not considered in this work. Here, Google Translator is also used to translate other language text to English and then classifies it into either non-suicidal or suicidal. All the audio files which are "transcribed automatically by using Google speech-to-text API" achieved an accuracy of about 95% in speech recognition tasks [6].

II. Methodology

Data

Data plays a very crucial role in this task. Here, the data used to classify suicidal tendencies is collected from Kaggle and this data is gathered from Reddit. With Exploratory data Analysis, there is no imbalance in data. The next step for this is to preprocess the data for modeling NLP techniques and then apply machine learning algorithms.

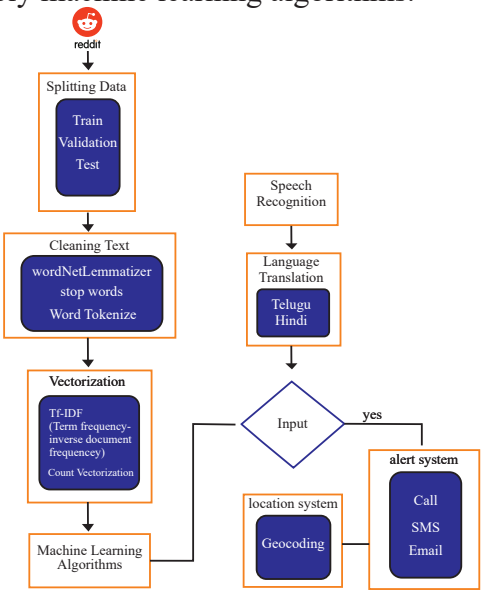


Fig. 1. *Methodology*

Data Preprocessing

This step is very important in any data mining or machine learning process that "involves transforming raw data into data in an understandable format" for NLP models. "Real-world data can be inconsistent, insufficient, and might contain many errors". This is the best method to solve the above-mentioned issues. This process helps in getting "better results through any machine learning algorithms". The two techniques that are also to be performed besides steps in data pre-processing are as follows:

1. *Word Lemmatization:* The goal of this process is "to reduce the inflectional forms of each word into a common base or a root word". Lemmatization is firmly connected to a process called stemming. The difference between the two can be said that "a stemmer usually operates on a single word without the knowledge of the context and hence it cannot separate between the words which have discriminant meanings relying on the part of speech. Stemmers are normally easier to implement and run" quicker.
2. *Stopwords:* "Stopwords are words in English that do not add meaning to a sentence". These words can be easily ignored as these words cannot forfeit the meaning of the sentence.
3. *Tokenization:* "This is a process of breaking a stream of text into words, phrases, symbols, or any other meaningful elements called tokens". This "list of tokens is considered as input for further processing. NLTK (Natural Language Toolkit) library has two functions like word_tokenize and sent_tokenize which can easily break a stream of text into a list of words or sentences respectively".

4. Vectorization Vectorization is a process where "a collection of text documents is turned into numerical feature vectors and this process is tokenization, counting, and normalization". This is also called a "Bag of words or Bag of n-grams"
5. *CountVectorizer:* Count Vectorizer implements "both tokenization and occurrence counting in a single class and can convert a collection of text documents to a matrix of token counts" [8].
6. *Tf – Idf Vectorizer:* "Tf-IDF stands for Term Frequency – Inverse document frequency" and it converts "a collection of raw documents into a matrix of Tf-IDf features. To re-weight count features into floating-point values which are suitable for usage by a classifier, a tf-idf transformer is used" [7].

III. Machine Learning Algorithms & Results

The data is trained in the below-mentioned algorithms and the accuracy of each algorithm is stated in the tabular concerning both CountVectorizer and TfIdfVectorizer. The best model is selected based on the accuracy of the model after evaluating through certain metrics like classification score which gives precision, recall, f1 score, and accuracy. In addition to this, an accuracy score and confusion matrix is also used. From the above table, the best model is the Support Vector Machine Classifier.

Using these Machine Learning algorithms, the text obtained is classified into suicidal and non-suicidal. Below are the observations and results that are obtained after using both Count Vectorizer and Tf-IDF vectorizer.

Cross-Validation Score

Algorithms	**Tf – IDF Vectorizer**	**CountVectorizer**
Logistic Regression	0.9112	0.9064

SVM Classifier	0.9157	0.8895
Gaussian Naïve Bayes	0.7488	0.7581
Multinomial Naïve Bayes	0.8293	0.8665
Decision tree	0.8407	0.8454
Random Forest	0.8903	0.8831
AdaBoost	0.8762	0.8708

Table 1. *Cross-Validation Score for different Vectorization processes*

Accuracy Score

Algorithms	**Tf – IDF Vectorizer**	**CountVectorizer**
Logistic Regression	0.9177	0.9089
SVM Classifier	0.9201	0.8933
Gaussian Naïve Bayes	0.7531	0.7635
Multinomial Naïve Bayes	0.8326	0.8600
Decision tree	0.8433	0.8414
Random Forest	0.8941	0.8834
AdaBoost	0.8819	0.8712

Table 2*. Accuracy Score for different Vectorization processes*

Confusion Matrix

Algorithms	**Tf – Idf Vectorizer**	**CountVectorizer**
Logistic Regression	[2623 212 265 2702]	[2619 216 312 2655]

SVM Classifier	[2607 228 235 2732]	[2555 280 339 2628]
Gaussian Naïve Bayes	[2235 597 835 2132]	[2351 484 888 2079]
Multinomial Naïve Bayes	[1913 922 49 2918]	[2099 736 76 2891]
Decision tree	[2509 326 583 2384]	[2490 345 575 2392]
Random Forest	[2527 308 306 2661]	[2461 374 302 2665]
AdaBoost	[2570 265 420 2547]	[2593 242 505 2462]

Table 3. *Confusion Matrix for different Vectorization Processes*

Classification Report

Algorithms	**Precision**	**Recall**	**F1 Score**
Logistic Regression	0.91 0.93	0.93 0.91	0.92 0.92
SVM Classifier	0.92 0.92	0.92 0.92	0.92 0.92
Gaussian Naïve Bayes	0.73 0.78	0.79 0.72	0.76 0.75
Multinomial Naïve Bayes	0.98 0.76	0.67 0.98	0.80 0.86
Decision tree	0.81 0.88	0.89 0.80	0.85 0.84

Random Forest	0.89 0.90	0.89 0.90	0.89 0.90
AdaBoost	0.86 0.91	0.91 0.86	0.88 0.88

Table 4*. Classification Report for Tf-IDF Vectorization*

Algorithms	**Precision**	**Recall**	**F1 Score**
Logistic Regression	0.89 0.92	0.92 0.89	0.91 0.91
SVM Classifier	0.88 0.90	0.90 0.89	0.89 0.89
Gaussian Naïve Bayes	0.73 0.81	0.83 0.70	0.77 0.75
Multinomial Naïve Bayes	0.97 0.80	0.74 0.97	0.84 0.88
Decision tree	0.81 0.87	0.88 0.81	0.84 0.84
Random Forest	0.89 0.88	0.87 0.90	0.88 0.89
AdaBoost	0.84 0.91	0.91 0.83	0.87 0.87

Table 5. Classification Report for CountVectorizer

IV. Design

Permissions

a. *Microphone:* This permission is needed to monitor the voice/speech of the user for identifying suicidal

tendencies. This voice or speech can be any conversation in a call or with a chatbot or searching in a browser.

b. *Keyboard:* This permission is needed to monitor the search in a browser, texting, etc..,.

c. *Location:* Location permission is very essential as the live location of the user is needed for the alert system. This permission cannot be reverted.

d. *Call:* Call permission is needed to call the emergency contact's mobile number from the user.

e. *Messages:* Messages permission is needed to send SMS for emergency contact's mobile number from the user.

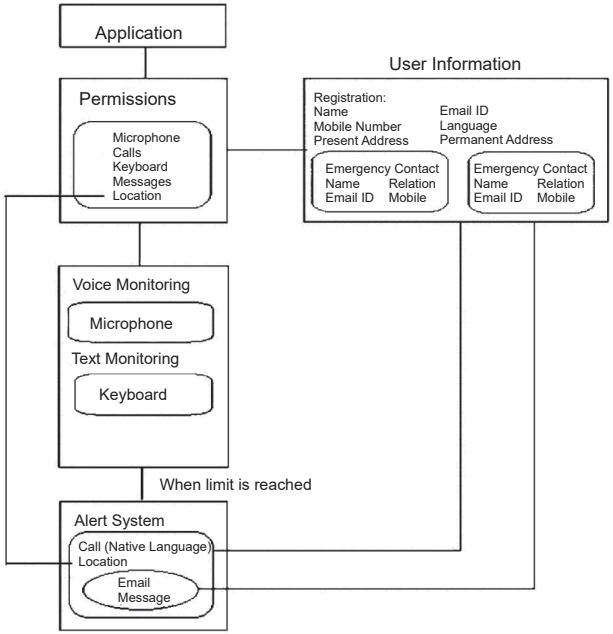


Fig 2. *Design*

User Information

Registration: User information is mandatory and it is the basic information of the registered user like

- Name
- Mobile Number
- Email address
- Language
- Address

Emergency Contact: This information is used to alert the emergency contacts provided by the user in the registration process. The number of Emergency contacts is as the user desires but a minimum of two contacts information is mandatory. Emergency contact information is as follows

1. Name
2. Mobile Number
3. Email Address
4. Relation
1. Parent/ Spouse/ Brother/ Sister/ Guardian/ Friend/ Cousin

Monitoring

1. *Voice Monitoring:*

   Voice/speech monitoring is done using a microphone which can be accessed through permissions. This works the same way as virtual assistant work [9]. Here, the user must register his voice after giving the user information and from then on, the microphone monitors the voice/speech. This can be done for calls/conversations with virtual assistants or through voice search. This voice is turned to text format and

then evaluation is done to predict suicidal or non-suicidal. This speech can be of any language and that can be translated using Google Translate and then the classification of text is done.

2. Text Monitoring:

   Text monitoring is done with keyboard permission. Language models are built to predict the most suitable word if we start typing and that is based on search history etc[9]. In the same way, the text is monitored and that can be either in browsing or texting (messaging). This text is used to classify whether the text is non-suicidal or suicidal.

Alert System

During voice/text monitoring, the text is classified and the counts of suicidal and non-suicidal are incremented whenever the text is classified into suicidal. A limit is set up so that if the count reaches the limit, an alert system is sent to the given contact information from the user. Alert system is done through call, SMS, and email where, both SMS and email have the link of the live location of the user which can be obtained from the location permissions earlier. In the call, the alert is made in the chosen language which is in User Information along with English and Hindi. In SMS, an alert message is given in the chosen language with the provided relationship and details of the user along with the link of the live location of the registered user. In the email, an alert message is sent along with the live location of the registered user with information in both languages.

V. Conclusion

By observing the accuracy score of various machine learning algorithms, it can be concluded that Support Vector Machine and Logistic regression give good results. But,

in testing, SVM proves to be the best model. This model predicts the translated sentences in an accurate way which is not satisfactory in Logistic Model. And that is because SVM has overfitting protection that will not necessarily depend on feature number since they have the ability in handling large feature spaces. SVM is best suited for problems containing "dense concepts and sparse instances". Better performance of SVM can also be observed due to its independence in choice of parameters which helps in overfitting protection. Theoretically, SVM acknowledges specific properties of text like "High dimensional feature spaces, less irrelevant features (dense concept vector), Sparse instance vectors". Due to "its ability to generalize good in high dimensional feature spaces, SVM eliminates the need for feature selection" which makes this algorithm best. Its robustness and avoiding any catastrophic failure and no requirement of parameter tunings like decision tree or random forest as it can find the best parameter automatically.[12]

VI. Future work

In the future, this design can be created as software or application which will be helpful for many individuals to surpass their depression and get help on time. This can also be implemented as a browser extension which helps to identify suicidal tendencies from browsing as well.

References

[1] https://www.who.int/news-room/fact-sheets/detail/suicide

[2] Of Veterans Affairs D, Others. National veteran suicide prevention annual report. Washington: Department of Veterans Affairs; 2019.

[3] Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. J Consult Clin Psychol. 1979;47:343–52

[4] Brown GK, Beck AT, Steer RA, Grisham JR. Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. J Consult Clin Psychol. 2000; 68: 371–7.

[5] Belouali, A., Gupta, S., Sourirajan, V. et al. Acoustic and language analysis of speech for suicidal ideation among US veterans. BioData Mining 14, 11 (2021). https://doi.org/10.1186/s13040-021-00245-y

[6] Cloud Speech-to-Text - Speech Recognition | Google Cloud [Internet]GoogleCloud.https://cloud.google.com/speech-to-text

[7] https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html#sklearn.feature_extraction.text.TfidfVectorizer

[8] https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html#sklearn.feature_extraction.text.CountVectorizer

[9] https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-language-model-nlp-python-code/

[10] https://www.verywellmind.com/why-do-people-commit-suicide-1067515

[11] https://medium.com/@bedigunjit/simple-guide-to-text-classification-nlp-using-svm-and-naive-bayes-with-python-421db3a72d34

[12] Li, Y., Bontcheva, K. and Cunningham, H., 2009. Adapting SVM for natural language learning: A case study involving information extraction. Natural Language Engineering, 15(2), pp.241-271.

Review on VGGNet: Architecture and Applications on Medical Sciences

Muralidhar Billa

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University, Bangalore, INDIA
mbillaMTECH2020@ced.alliance.edu.in

Keerthivarman G

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University Bangalore, INDIA
gkeerthivarmanMTECH2020@ced.alliance.edu.in

Dr. Shobana Padmanabhan

Associate Professor – Dept. of Information Technology,
Alliance College of Engineering and Design
Alliance University, Bangalore, INDIA
shobana.padmanabhan@alliance.edu.in

***Abstract*—**VGGNet is also a subset of a convolutional neural network (CNN). Mainly The concept of VGGNet is to productively give the boost for the model that is performed.

Previously, deep learning has proved its best in computer vision tasks. As new technology is growing up with machine learning rapidly. Deep learning overcomes the traditional way of "machine learning which enables the new learning applications in computer vision, medical industry, Speech recognition, machine translation", image recognition, etc. In this paper, VGGNet mainly concentrated on how it performs efficiently in the medical field and reviewed. The VGGNet has 16 and 19 layers. By using those layers, the model can be reusable to the data that was captured by the information, and the feature size is minimized, and the accuracy is increased. The structure and implementation of the model give the improved accuracy that can be deployed into the CNN which is related to VGGNet.

Keywords—*VGGNet 16, VGGNet 19, Architecture.*

I. Introduction

VGGNet is known as Visual Geometry Group Network, it is the neural network that helps CNN with multiple layers. A simple VGGNet can start with 3 layers to 19 layers of VGG mostly these many layers are used to identify the object and detect with accurate measures. VGG 16 and VGG 19 are weight layers those layers enhance the classification accuracy by increasing the depth of a convolutional neural network. However, this VGGNet is for Object Detection, it cannot be able to detect the scenes recognition. The latest version of VGGNet can be able to read the scenes. In general, there are five max-pooling channels implanted between the layers the portrayal of info was like a picture, hidden-layer output matrix, and so on. VGGNet takes the absolute of 224*224 RGB pictures and goes through convolutional layers which channel the picture.

Assume that VGGNet is trained with some dataset that consists of 1.5 million preparation pictures, 900,000 testing

pictures, and 60,000 approval pictures. That model has been gotten some precision 94% test exactness by Image net. VGGNet has been the best real-world application in many aspects mostly in medical such as heart rate based on brain detection, chest x-ray detection, and motion of the human body. The main contribution of VGGNet is that stacking multiple filters without increasing the pool could gradually raise the depth of networks with the help of minimizing the total number of parameters. Where VGGNet is much more overpriced than the other models behind that it also uses a lot more memory. Layers can be removed by revealing parameters that are connected to layers reducing the necessary parameters. VGGNet can work with 16 and 19 layers. VGGNet 16 has a test accuracy of 92.7% in general, whereas for 19 layers 3 more convolutional layers can give the model more accuracy and the VGGNet 19 has the same concept as what 16 layers had.

II. ARCHITECTURE

VGGNet is based on features of CNN there are as follows in the below figure:

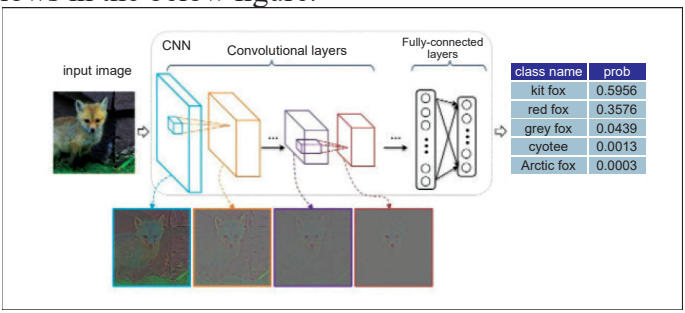


Fig. 1. *Architecture of CNN*

CNN is a supervised learning method that includes input, hidden and an output layer. Coming to the hidden layer there are multiple layers, inside all the detection is done. Where in total three layers can divide into Convolutional, Pooling,

and fully connected (FC). CNN's main motive is to draw the necessary content from the data that contains mirror reflection which centers the system to be assembled that most precisely fits with the type of data. The engineering of CNN is from an info picture which is regularly having the neighbor associations and loads are restricted with the surveying which causes interpretation changes in the information. A CNN comprises at least one completely associated layer that is introduced in convolutions and pooling layers. The principal design of the CNN network is tending to further develop exactness by input the layers of various models like GoogleNet, VGGNet, LEENet, and so on Where each strategy of the model is the same yet has a unique type of approach that gives an efficient model that must provide the better accuracy of the given data. CNN is one of the best methods to use in computer visions, Image processing, and other uses.

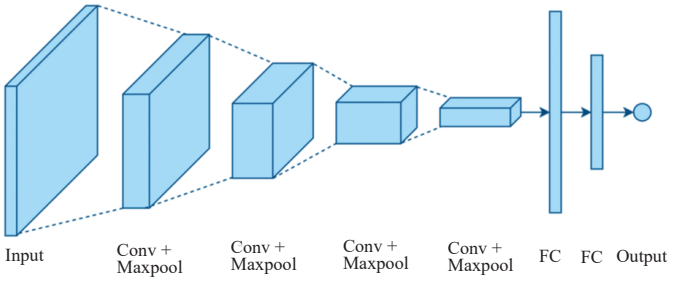


Fig. 2*: A basic architecture of CNN*

1. Convolutional Layer:

As the name suggests, the convolution layer plays a huge capacity in how CNN works. It frames the basic unit of a ConvNet. whenever data strikes a convolution layer, the layer itself evolves to every one of the channels across the spatial dimensionality of the data to give a 2D actuation map. The result of neurons that can relate to neighborhood districts of the info might be laid out utilizing the

convolution layer. Neurons that comprise indistinguishable trademark maps stock the heap (boundary sharing) consequently diminishing the intricacy of the organization with the guide of saving the wide assortment of boundaries low. The corrected straight unit (ordinarily abbreviated to ReLU) targets to utilize an 'element wise' enactment work like sigmoid to the result of the actuation made by utilizing the past layer. Convolution layers can broadly downsize the intricacy of the adaptation through the improvement of its result.

2. Non-linearity layer:

This layer applies to the non-soaking enactment work. It will blast the nonlinear places of the decision trademark and of the general local area which are ideally suited for multi-facet networks while not influencing the responsive fields of the convolution layer. The actuation work is typically sigmoid, tanh and ReLU. Contrasted with unmistakable capacities redressed Linear Units are wanted. Likewise, to improve the general exhibition of the organization, the SOFTMAX enactment trademark is employed on the top of the last layer.

3. Pooling Layer:

"CNN conveys not exclusively convolution layers yet additionally, conjointly some pooling layers. There can be a pooling layer in a flash after a convolutional layer. It proposes that the results of the convolution layers are the contributions to the pooling layers of the local area. Pooling activities cut the size of the trademark maps through the exploitation of certain elements, to sum up, subregions, such as taking the normal or the greatest cost. The pooling layer's objective to grade by grade diminish the dimensionality of information, and thusly additionally reduce the number of boundaries and the interaction intricacy of the model and

therefore control the issue of overfitting. A few the normal pooling tasks are max pooling, normal pooling, stochastic pooling, unearthly pooling, spatial pyramid pooling, L2-standard pooling, and multiscale request significantly less".

There are two regularly resolved procedures of max pooling. That Typically. Besides, covering pooling is in like manner applied, wherein the step is 2 with a bit longer. In any case, because of the harmful idea of pooling, having a piece length better than 3 can sometimes apparently bring down the presentation of the model. It is similarly crucial to realize that other than max-pooling, CNN models should contain the by and large pooling. General pooling layers are comprehensive of pooling neurons which may be equipped to play out a huge assortment of ordinary activities at the same time with L1/L2-standardization, and normal pooling.

4. Fully connected layer:

The high stage thinking inside the brain's local area is done through totally connected layers. In a completely related layer Neurons at a couple of stages include associations with every one of the actions inside the first layer, as seen inside the traditional Multilayer Perceptron (MLP) brain organizations. Their actions will appropriately be processed with a framework activity followed by utilizing an inclination offset. The result of the convolution and pooling layers, adding a completely related layer is, in addition, a less expensive technique to break down the non-straight blends of these elements.

However, the limit of abilities is the convolutional and pooling layers. It may be adequate for the class task; however, combinations of these elements might be surprised the best. "The neurons in a completely related layer are not spatially arranged consequently there can't be a convolution layer after a completely associated layer. A related layer passes

the two-layered result to the result layer any place we can utilize a SoftMax highlight or a sigmoid to anticipate the enter class name". As of late, a couple of designs supplanted their FC layer. Be that as it may, the target of the associated layer is to smooth the exorbitant stage of the elements, which can be learned via the convolutional layers, and to mix every one of the variables.

In convolutional Neural organization presently take a gander at the VGG engineering and how it performs:

Contribution:

In VGGNet takes in the picture input size length of the photo is predictable which can estimate the accuracy of the FC.

Convolutional Layers: "ReLU stands for rectified linear unit activation function; it's miles a piecewise linear characteristic with a purpose to output the enter if high quality; in any other case, the output is zero. The convolution stride is fixed at 1 pixel to maintain the spatial decision preserved after convolution (stride is the variety of pixel shifts over the input matrix).

Hidden Layers: All the hidden layers within the VGG community use ReLU. VGG does no longer typically leverage Local Response Normalization (LRN) as it increases reminiscence consumption and schooling time. Moreover, it makes no improvements to normal accuracy.

Fully Connected Layers: The VGGNet has three related layers. Out of the three layers, the first have 4096 channels every, and the 1/3 has a thousand channels, 1 for every class".

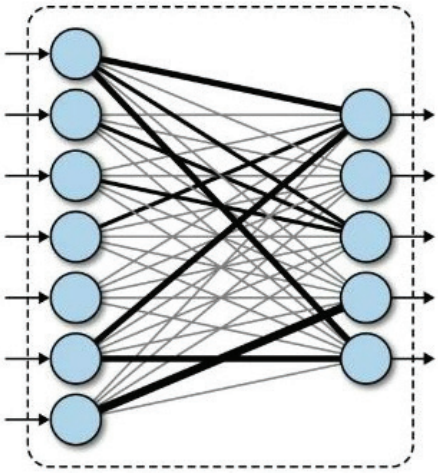


Fig. 3*: Fully Connected Layers*

VGG 16 Architecture:

The assortment 16 in the call VGG alludes to the reality that it is 16 layers of profound brain organization (VGGNet). This implies that VGG16 is a very enormous local area and has a total of around 138 million boundaries. In any event, as indicated by popular principles, it's far a major local area. Notwithstanding, VGGNet16 engineering's effortlessness makes the local area extra alluring. Just by taking a gander at its construction, it tends to be expressed which is in an equal state. By taking a gander at the number of channels that we will use, around 64 channels are accessible that we will twofold to roughly 128 after which to 256 channels. In a definitive layer, we will utilize 512 channels.

The intricacy and requesting circumstances

The number of channels that can utilize copies on each progression or through each heap of the convolution layer. This is an excellent statute used to format the engineering of the VGG16 organization. And that implies that it requires

some investment to prepare its boundaries. “Given its force and wide assortment of completely connected layers, the VGG16 variant is more than 533MB. This makes executing a VGG network a period-eating task.

The VGG16 adaptation is used in various profound learning photograph type inconveniences, but more modest local area designs along with GoogleNet and SqueezeNet are consistently generally suitable. Regardless, the VGGNet is a phenomenal structure block for acquiring information on capacities since it is not difficult to place into impact”.

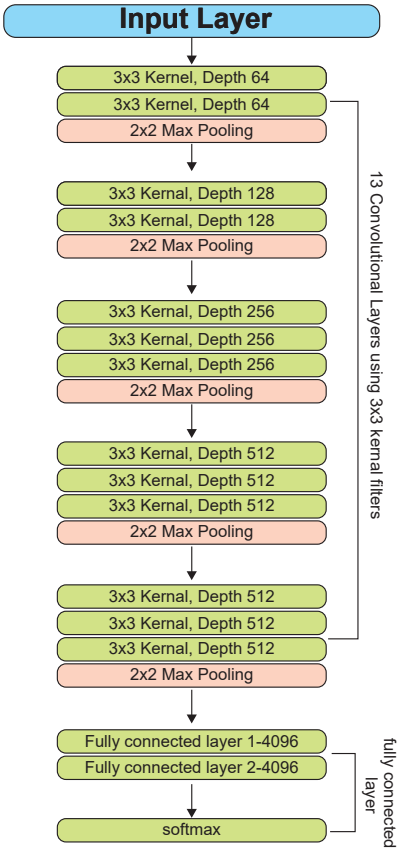


Fig. 4: *VGG-16 Architecture of a VGG16 model*

Execution of VGG Models

"VGG16 very outperforms the former adaptations of models inside the ILSVRC-2012 and ILSVRC-2013 contests. Besides, the VGG16 result is seeking the kind of venture champ (GoogleNet with 6.7% blunder) and broadly beats the ILSVRC-2013 prevailing accommodation Clarified. It got eleven.2% with outside preparing measurements and around 11.7% without it. As far as the unmarried-net execution, the VGGNet-sixteen model accomplishes the great outcome with roughly 7.0% gander at blunders, in this manner outperforming an unmarried GoogleNet through around 0.9%".

VGG 19**:** This model idea is pretty much as same as VGG 16 with additional 3 layers added to it. In complete, there are 16 convolutional, 3 completely associated layers, 5 Max pool layers, and 1 SoftMax layer. The figure is as per the following

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 x 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **covn3-64**	conv3-64 **covn3-64**	conv3-64 **covn3-64**	conv3-64 **covn3-64**
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 **conv3-256** **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512** **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512** **conv3-512**
maxpool					
FC - 4096					
FC - 4096					
FC - 1000					
soft-max					

Fig. 5*: VGG 19 model*

To diminish the scope of boundaries in such profound organizations, "it utilizes little 3×3 channels in all convolutional layers and decent used with its 7.3% goofs charge. The VGG-19 model changed into now not the victor of ILSVRC30 2014, nonetheless, the VGG Net is one of the most persuasive papers as it fortified the conviction that CNNs should have a profound organization of layers for this progressive outline of noticeable data to work. Keep it profound. This model is taught on a subset of the ImageNet27 information base, which is utilized inside the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC).30 The VGG-19 is prepared on more than 1,000,000 photographs and may group pictures into 1,000 thing classes, for instance, console, mouse, pencil, and bunches of creatures. Thus, the model has figured out rich capacity that portrayals for a broad scope of pix.

Locale-based CNN (R-CNN)31 creates a region idea and utilizations CNN for thing recognition and classification. Quick R-CNN32 and Fasted R-CNN33 are proposed later to pick up the pace of the framework and work on the precision". In any case, the R-CNN strategy isn't generally reasonable for bosom disease identification since mammograms significantly shift in surface and length of sore (most tumors) from one case to another.

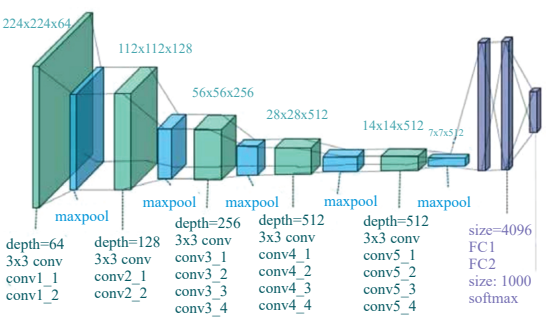


Fig 6: *Architecture of VGG-19 Model*

III. APPLICATIONS USING VGGNet

Cerebrum Tumor Classification:

Cerebrum cancer therapy is the exceptionally trickiest errand to fix in clinical history without having any deficiency of the human. In 2014 VGGNet has positioned the second situation for picture investigation and giving the precision of mistake at its ideal. Where 16-19 layers are been utilized for picture handling the top blunder rate for this is 6.8%. The assignments for CNN-based engineering are grouping, limitation, Dataset here utilized considering ImageNet and the execution climate is four NVIDIA, titan Black, GPU. Here the grouping mistake decreases with a steady increment of the profundity and soaked when the profundity is reached at 19 layers. In this Paper, characterization execution by utilizing various structures VGG-16 performed 98.06%,98.14%,98.71% for having different intervals0f age of 25,50,90 where VGG16 is consuming a long preparation time when contrasted with any remaining models due to more profound layers. At last, this paper addresses there are benefits and inconveniences of involving VGG in Brain Tumor Classification because, even though it gives better precision, it likewise consumes time.

Melanoma Detection:

It is the most hazardous skin disease, which should decrease in the future. To stay away from this kind of disease in this paper, by thinking about the great quality pictures and doing the preprocessing steps that can capably distinguish the strategies to the class skewness issue. Profound brain organizations. While contrasted with VGG 16 and 19, 19 is more costly than 16 because the layer that introduced in that so the quantity of preparing and test pictures has been taken in when the increase is as per the following

	BCC	BKL	MEL	NV	VASC	DF	AKIEC
Original Training Data	247	815	801	4975	109	85	253
Augmented Training Data	4773	3275	3308	1111	3488	4152	3856
Original Test Data	139	284	312	1730	33	30	74
Augmented Test Data	1692	1479	1664	1111	1484	1632	1552

Fig. 7: *Data considered*

The best values that have been predicted by the output for all the architecture is as follows:

Methods	Precision	Recall	F-Score	ACC
AlexNet	0.7717	0.7873	0.7726	0.7853
AlexNet*	0.8421	0.8125	0.8231	0.8045
AlexNet50	0.8652	0.8663	0.8537	0.8637
ResNet50*	0.9373	0.9253	0.9274	0.9208
VGGNet16	0.8442	0.8447	0.8433	0.8436
VGGNet16*	0.907	0.9032	0.9061	0.8836
VGGNet19	0.8468	0.8457	0.8436	0.8461
VGGNet19*	0.8855	0.8882	0.8838	0.8870
Xception	0.4472	0.6633	0.5346	0.6629
Xception*	0.9019	0.9057	0.9041	0.9030

Fig. 8*: CNN classification models*

After increase and preprocessing, has ninety-two% precision while "AlexNet and VGG19 have exactness eighty% and 88% individually. This implies the space in precision is 12% and 4% in want of ResNet50in correlation with the AlexNet and VGG19. The opening of ResNet50 contrasted with Xception and VGGNet16 is 26% and eight% individually". Along these lines, ResNet50 has superb precision and the Xception has the most horrendously terrible exactness among all partners. Comparative ends will be withdrawn from another class proportion. "Assuming the general execution of ResNet50 adaptation is contrasted with the agreeing with the execution of ResNet50 and VGG16, then the deviation in precision is 6% and 4% individually", and in look at with Xception structure Improved by 20%.

Methods	AKIEC	BCC	BKL	DF	MEL	NV	VASC	Average ± Std
AlexNet	0.9111	0.8492	0.7053	0.9836	0.80	0.6816	0.9605	0.84 ± .11
VGGNet16	0.9726	0.8826	0.7535	0.9933	0.9438	0.7871	**1**	0.90 ± 0.09
VGGNet19	0.8749	**0.961**	**0.8262**	1	0.8764	0.7625	**1**	0.89 ± 0.08
ResNet50	**0.9743**	0.929	0.7717	0.9905	**0.9465**	**0.9232**	**1**	**0.92** ± 0.07
Xception	0.9646	0.914	0.7652	1	0.8911	0.8427	1	0.90 ± 0.08

Fig. 9*: Classification performance on different classes*

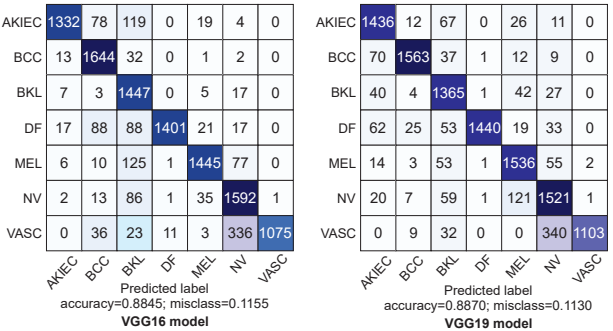


Fig. 10: *Confusion matrix of VGG models*

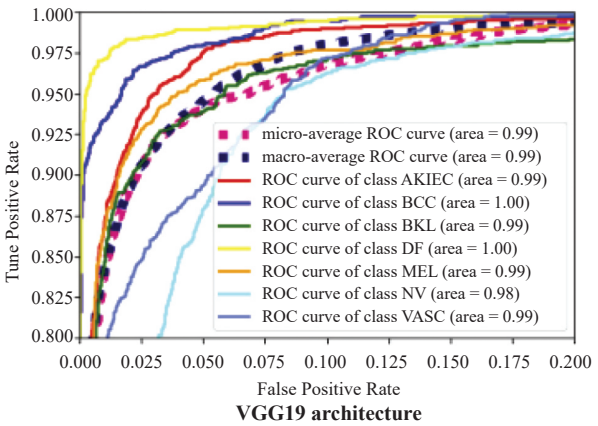


Fig. 11*: Roc Curve of VGG 19*

Breast cancer identification and classification on VGG:

This gives a nitty-gritty examination of pre-prepared CNN designs and moves to learn. Likewise, talk about the exploratory outcomes acquired after applying the proposed

approach alongside its exhibition assessment. “VGGNet is similar as AlexNet other than with additional convolution layers. VGGNet comprises of 13 convolutions, amendment, pooling and 3 completely connected layers”. The conv local area utilizes 3 × 3 windows size sift through and a couple of × 2 pooling local area. VGGNet plays higher when contrasted with AlexNet due to its basic design.

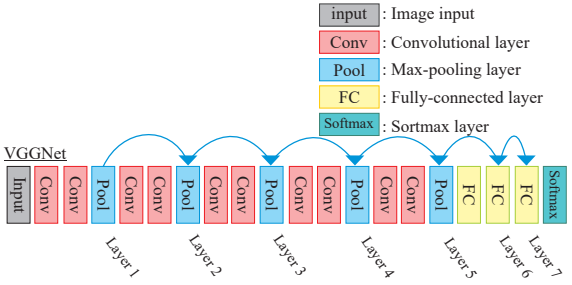


Fig. 12: *Architecture of VGGNet*

As shown in fig 12 the magnification and comparative analysis of CNN architecture for VGG is 94.15%.which is best but other architectures can also beat VGG

CNN Architectures	Lens	Magnification			Average Classification Accuracies
	100X	140X	200X	500X	
GoogLeNet	90.4	93.7	95.3	94.6	93.5%
VGGNet	90.8	94.8	96.7	94.2	94.15%
ResNet	91.5	93.3	95.4	97.2	94.35%
Proposed Framework	96.8	96.9	97.8	98.6	97.525%

Fig. 13: *Magnification and Accuracies*

Chest X-ray images for Covid -19

This paper, The COVIDX-Net comprises seven explicit structures of the profound convolutional brain, local area models, which incorporate altered both VGG19 and Google MobileNet with two dimensional. Every profound neural link model is in a situation that could break down the standardized powers, X-beam picture will arrange the patient’s ability whether it is good or bad with symptoms of covid. strongly accomplishment dependent absolutely upon

eighty-20% of X-beam pics for the adaptation instruction and evaluating stages, individually. The VGG19 and DenseNet models affirmed the marvelous with similarly with stand-by and large execution of mechanized covid sort with f1-scores of zero.89 and 0.91 for standard and COVID-19, individually. “The most exceedingly terrible class generally execution is obtained for the InceptionV3 model with f1-rankings of 0.67 for ordinary cases and 0.00 for COVID-19 cases”.

The upsides of execution measurements of every profound getting to realize classifier are given in Table three. The greatest accuracy of profound learning classifier anyway their comparing exhibitions had been most horrendously terrible to group the ordinary cases. Along these lines, by encouraging the VGG19 and DenseNet201 styles that are to be applied for inside the constructions that recognize their wellbeing ubiquity of patients prepared profound concentrating on classifiers with exactness and move-entropy (misfortune) inside the preparation and approval step. The excellent appraisals of tutoring and approval exactness had been achieved for VGG19 and DenseNet201.

Classifier	Patient Status	Precision	Recall	F1- score
VGG19	*COVID-19*	0.83	1.00	0.91
	Normal	1.00	0.80	0.89
DenseMet201	*COVID-19*	0.83	1.00	0.91
	Normal	1.00	8.00	0.89
ResNetV2	*COVID-19*	1.00	0.40	0.57
	Normal	0.62	1.00	0.77
InceptionV3	*COVID-19*	0.00	0.00	0.00
	Normal	0.50	1.00	0.67
InceptionResNetV2	*COVID-19*	1.00	0.60	0.75
	Normal	0.71	1.00	0.83
Xception	*COVID-19*	1.00	0.60	0.75
	Normal	0.71	1.00	0.83
MobileNetV2	*COVID-19*	1.00	0.20	0.33
	Normal	0.56	1.00	0.71

Fig. 14*: Performance of VGG with other Models*

VGGNet performs the highest F-1 score which implies the performance of the model concerning the deeper layers of images analysis in covid-19 x-ray images VGG gives their best.

Lung cancer Identification:

The significant standard of the past couple of layers of the VGG-16 adaptation had been changed to oblige new picture classifications in the current utility. The pre-instructed model with switch acquiring information on containing normal notwithstanding uncommon MR previews with selective neurological illnesses.

The realities set changed into then apportioned the utilization of a 10-crease move-approval instrument. The approval on the gander at set the utilization of awareness of execution in assessment to the contrary present status of-the-craftsmanship. Besides, the methodology gives order a finish to-stop shape on crude pictures with no hand-crafted property extraction.

Pre-instructed on a tremendous, characterized homegrown photo dataset like ImageNet has ended up being useful for picture arrangement inconveniences through move acquiring information. In DCNN, portrayals are assimilated at various organization layers connected. “The design of VGG_16 As characteristic the conv1 layer gets an image with a length equivalent to 224 x 224 because the enter”. Along these lines, information photograph is engendered using a fixed of Conv layers having an open subject of 3×three. “The convolution step is taken as 1 pixel. Five max-pooling layers with step equivalent to 2 are taken for spatial pooling (down inspecting)”. Max-pooling layers are victorious a few of the conv layers and their miles are completed over a 2×2-pixel window. Layer, each neuron

acknowledges enter from the enactments of the past layer neuron.

Ref	Methodology	Dataset	Features	Se	Sp	Acc	Partitioning scheme
[11]	DWT+PCA+ PSO optimized Kernel SVM	90 images (SH, 85 Ab)	1024	98.12	92	97.78	S Fold CV
[12]	DWT+BBO optimized SVM	5 normal and 85 abnormal	10	98.12	92.00	97.78	5x5 fold CV
[13]	DWT+BAT optimized Learning Machine	18 normal and 114 abnormal	7	99.04	93.89	98.33	10x10 fold CV
[27]	Ripplet-II Features+ PCA+Modified PSO based ELM	20 normal and 140 abnormal	13	99.64	99.50	99.62	5 -fold CV
[28]	Pseudo Zernike moment+ Kernel SVM	20 normal and 140 abnormal	400	99.93	98.50	99.75	10x5-fold
Proposed	VGG-16 with transfer learning	20 normal and 140 abnormal	NA	100	100	100	10-Fold CV

Se: Sensitivity; *Sp:* Specificity; *Acc:* Accuracy
BBO: Biogeography based optimization
CV: Cross Validation
DWT: Discrete Wavelet Transform
ELM: Extreme Learning Machine
PCA: Principal Component Analysis
PSO: Particle swarm optimization
SVM: Support Vector Machine

Fig. 15*: comparing the VGG-16 pre-trained, the transferred learning.*

Pre-taught VGG-16 DCNN rendition with move acquiring information has been examined for the odd MR mind photo class. This VGG-16 model with move concentrates on becoming progress in offering a one hundred% notoriety expense. The results have been progressed to the predominant old-style procedures articulated in the mind picture class.

VI. Conclusion

In the wake of considering the audits of the multitude of papers, VGGNet is best for the mental pictures, cellular breakdown in the lung's recognition, skin malignant growth location, etc. Since VGG 16 and 19 both are doing their model exhibition yet utilizing 19 is pricey when contrasted with VGG 16. Normally, this sort of model is utilized for detailed picture investigation without having any blunder in the clinical field. The presentation of this model will be over 95-100 percent due to its numerous layers and convolutional, pooling and SoftMax work. For fewer Epochs, VGG comes to 100 percent. VGG given CNN shows the possibility to play out the picture investigation in a cellular breakdown in the lungs identification, covid19 chest x-beam, and so on more viable grouping should be finished. As contrasted and a few Architectures with VGGNet it gives 4%-5% less and other than those it plays out awesome. VGGNet is great disregarding the expense sometimes.

References

[1] VGG16 – Convolutional Network for Classification and Detection," 20 November 2018. [Online]. Available: https://neurohive.io/en/popular-networks/vgg16/. [Accessed 2019 April 30].

[2] H. Lee and M. Whang, "Heart Rate Estimated from Body Movements at Six Degrees of Freedom by Convolutional Neural Networks," Sensors, vol. 18, pp. 1-19, 2018.

[3] American College of Radiology, ACR BI-RADS — Mammography, Ultrasound & Magnetic Resonance Imaging (4th ed.), American College of Radiology, Reston, VA (2003).

[4] Kegelmeyer, Jr., W.P., Pruneda, J.M., Bourland, P.D., et al., "Computer-aided mammographic screening for

speculated lesions," Radiology, vol. 191, pp.331–337 (1994).

[5] Liu, S.L., Babbs, C.F., and Delp, E.J., "MultiResolution Detection of spiculated Lesions in Digital Mammograms," IEEE Transactions on Image Processing, vol. 10, pp.874-884 (2001).

[6] Matsubara, T., Fujita, H., Endo, T., Horita, K., et al., "Development of mass detection algorithm based on adaptive thresholding technique in digital mammograms," presented at Digital Mammogrpahy (1996).

[7] Chelghoum, R., Ikhlef, A., Hameurlaine, A., Jacquir, S. (2020). Transfer Learning Using Convolutional Neural Network Architectures for Brain Tumor Classification from MRI Images. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds) Artificial Intelligence Applications and Innovations. AIAI 2020. IFIP Advances in Information and Communication Technology, vol 583. Springer, Cham. https://doi.org/10.1007/978-3-030-49161-1_17

[8] "American Cancer Society," 2018. [Online]. Available: https://www.cancer.org/cancer/ melanoma-skin-cancer/about/key-statistics.html. Amoroso, N., et al., 2018.

[9] Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge. J. Neurosci. Methods 302, 3–9. Ayan, E.H.M.Ü., 2018.

[10] Data augmentation is important for the classification of skin lesions via deep learning. Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT). Barata, C., Marques, J.S., Celebi, M.E., 2014.

[11] Improving dermoscopy image analysis using color constancy. Image Processing (ICIP), 2014 IEEE International Conference on. pp. 3527–3531.

[12] SanaUllah Khana , Naveed Islama , Zahoor Jana , Ikram Ud Dinb , Joel J. P. C Rodrigues c,d,e,∗ a Islamia Col-

lege University Peshawar, Pakistan b The University of Haripur, Pakistan c National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, MG, Brazil d Instituto de Telecomunicações, Portugal e Federal University of Piauí (UFPI), Teresina, PI, Brazil

[13] Ezz El-Din Hemdan1, Marwa A. Shouman1, and Mohamed Esmail Karar 2,3, IEEE, Member 1Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Egypt. 2Department of Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, Egypt. 3Department of Computer Engineering, College of Computing and Information Technology, Shaqra University, Saudi Arabia

[14] Shailesh Kumar Thakur1 & Dhirendra Pratap Singh1 & Jaytrilok Choudhary1 # Springer Science+Business Media, LLC, part of Springer Nature 2020

[15] Taranjit Kaur1, Tapan Kumar Gandhi2 Department of Electrical Engineering Indian Institute of Technology (IIT Delhi), Hauz Khas, New Delhi-110016, INDIA Email: taran.rehal@yahoo.com1; Taranjit.Kaur@ee.iitd.ac.in1, tgandhi@ee.iitd.ac.in2

[16] M. C. Lee and S. J. Nelson, “Supervised pattern recognition for the prediction of contrast-enhancement appearance in brain tumors

[17] from multivariate magnetic resonance imaging and spectroscopy,”

[18] Artif. Intell. Med., vol. 43, no. 1, pp. 61–74, 2008.

[19] E. I. Zacharaki, S. Wang, S. Chawla, and D. Solo, “Classification

[20] of brain tumor type and grade using MRI texture and shape in a

[21] machine learning scheme,” Magn. Reson. Med., vol. 62, no. 6, pp.

[22] 1609–1618, 2009.

[23] S. Wang, S. Kim, S. Chawla, R. L. Wolf, W. Zhang, D. M. O. Rourke, K. D. Judy, E. R. Melhem, and H. Poptani, "Differentiation between Glioblastomas and Solitary Brain Metastases Using Diffusion Tensor Imaging," Neuroimage, vol. 44, no. 3, pp. 653–660, 2010.

[24] D. J. Hemanth, C. K. S. Vijila, A. I. Selvakumar, and J. Anitha, "Performance Enhanced Hybrid Kohonen-Hopfield Neural Network for Abnormal Brain Image Classification," in Signal Processing, Image Processing, and Pattern Recognition, Springer, 2011, pp. 356–365.

[25] F. G. Zollner, K. E. Emblem, L. R. Schad, F. G. Zöllner, K. E. Emblem, and L. R. Schad, "SVM-based glioma grading: Optimization by feature reduction analysis," J. Med. Phys., vol. 22, no. 3, pp. 205–214, 2012.

[26] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Segmentation, feature extraction, and multiclass brain tumor classification," J. Digit. Imaging, vol. 26, no. 6, pp. 1141–1150, 2013.

[27] K. Skogen, A. Schulz, J. B. Dormagen, B. Ganeshan, E. Helseth, and A. Server, "Diagnostic performance of texture analysis on MRI in grading cerebral gliomas," Eur. J. Radiol., vol. 85, no. 4, pp. 824–829, 2016.

[28] N. B. Bahadur, A. K. Ray, and H. P. Thethi, "Comparative Approach of MRI-Based Brain Tumor Segmentation and Classification Using Genetic Algorithm," J. Digit. Imaging, vol. 31, pp. 477–489, 2018.

[29] S. Lahiri, "Glioma detection based on multi-fractal features of segmented brain MRI by particle swarm optimization techniques," Biomed. Signal Process. Control, vol. 31, pp. 148–155, 2017.

[30] N. Gupta, P. Battle, and P. Khanna, "Glioma detection on brain MRIs using texture and morphological fea-

tures with ensemble learning," Biomed. Signal Process. Control, vol. 47, pp. 115–125, 2019.

[31] Y. Zhang, S. Wang, G. Ji, and Z. Dong, "An MR brain images classifier system via particle swarm optimization and kernel support vector machine," Sci. World J., vol. 2013, 2013.

[32] G. Yang, Y. Zhang, J. Yang, G. Ji, Z. Dong, S. Wang, C. Feng, and Q. Wang, "Automated classification of brain images using wavelet-energy and biogeography-based optimization," Multimed. Tools Appl., vol. 75, pp. 15601–15617, 2016.

[33] https://www.bing.com/search?q=vgg+19+architecture&cvid=f35fda919fe249bda4c2b040b8485719&aqs=edge.1.69i59l3j0l3j69i60j69i61j69i60.3908j0j1&pglt=41&FORM=ANSAB1&PC=U531&ntref

An Overview of LeNet: Architecture and Applications in Medical Sciences

Shanti Konda

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University, Bangalore, INDIA
kshantiMTECH2020@ced.alliance.edu.in

Pankaj Kumar Singh

Student, Computer Science, and Engineering
Alliance College of Engineering and Design
Alliance University, Bangalore, INDIA
kpankajMTECH2020@ced.alliance.edu.in

Dr. Shobana Padmanabhan

Associate Professor – Dept. of Information Technology
Alliance College of Engineering and Design
Alliance University, Bangalore, INDI
shobana.padmanabhan@alliance.edu.in

***Abstract*—**Among published CNNs in the performance of computer vision tasks, LeNet is the first to capture wide attention. This is done by successfully training CNNs

with backpropagation. LeNet is proved to be a dominant approach in supervised learning by achieving matching results with a Support vector machine. LeNet is initially used to recognize digits for processing deposits in many ATMs which can be observed even these days. LeNet architecture is used to read zip codes, digits, etc., and can process images with higher resolution as it needs large firmer layers. LeNet has been evolved over the past 2 decades and there are different LeNet architectures. The applications of LeNet-5 architecture are mainly observed in medical image processing, analysis, and some disease detection.

Keywords— *LeNet, Deep Learning, Architecture, medical image processing, Detection, Convolutional Neural Network (CNN).*

I. Introduction

LeNet is one of the Convolutional Neural networks (CNN) architecture proposed by Yann LeCun et al. in the year 1989. LeNet is usually referred to as LeNet – 5, a simple CNN which is a type of feed-forward neural network where artificial neurons respond to a part of neighboring cells in the range of its converge, and, these networks perform well with large-scale images in image processing. This Lenet-5 can be said as one of the earliest CNN that helped in promoting the development of Deep Learning [17].

In 1989, Yann LeCun et al. for some practical applications, a backpropagation algorithm is applied. LeNet firstly is a combined CNN that is trained by using backpropagation algorithms for reading handwritten digits or numbers and LeNet is then successfully applied in recognizing handwritten zip code numbers which are provided by US Postal Service [17].

LeNet architecture can be said as a straightforward and small architecture in terms of memory and makes perfect

in teaching the basics of CNNs by running on CPU if the system does not have a suitable GPU hence, LeNet is a great 'First CNN' [16]. In this paper, there is a review of the fundamental architecture of LeNet and its variations according to LeCun and a description of several applications of LeNet in medical sciences

II. Architecture

LeNet is a group of CNNs where the networks are broadly considered to be the first set of true CNNs. These neural networks are capable of classifying images of small single-channel which are in black and white and that too with promising results [19]. LeNet consists of three different networks and they are as follows:

- LeNet – 1
- LeNet – 4
- LeNet – 5

Before entering the LeNet description, the basics of CNNs are important and that is as follows.

Convolutional Neural Network basics

CNNs are initially designed to analyze visual imagery or to imitate the biological process of "human vision". In the architecture of every CNN, three layers are observed and those layers can be observed along with the typical view of a CNN architecture in Figure 1 below.

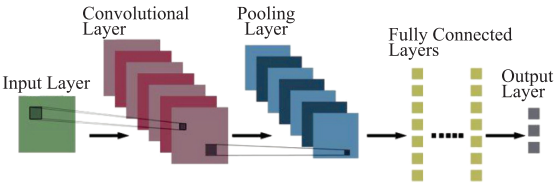


Figure 1. *The basic architecture of CNN*

Convolutional Layer

The convolutional layer plays a dominant role in the operation of CNN. It is a fundamental part of a CNN where computation takes place. The parameters of this layer mostly concentrate on the usage of learnable kernels, or it can be said that the two – dimensional (2D) convolutional layer uses trainable kernels or filters for performing convolution operations, like trainable bias which is optional for each kernel. These learnable kernels are small in spatial dimensionality but unfold the whole dimension of the input's depth and the layer "convolves each kernel across that spatial dimensionality to give 2D activation map" [11] when the information hit the convolution layer. These operations involve shifting the "trainable kernels or filters over the input in step process" which can be called strides. Usually, if "the stride is huge, more spaces can be ignored by kernels between every convolution.

This gives lesser convolutions and a more small-scale output size. A multiplication operation is implemented between input and the kernel with a bias term and that is added to the result and this operation is seen for each placement of a given kernel". This produces a feature map with the convolved result, and these maps are "gone through activation function" which provides input for the next layer. The size of the feature map can be obtained by the following expression.

$$\frac{(input_{size} - kernel_{size} + 2 * padding)}{stride} + 1$$

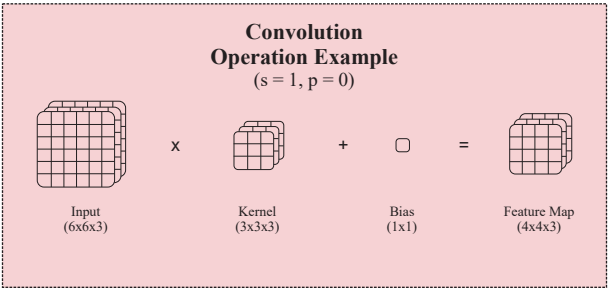


Figure 2. *Convolution operation example*

Convolution operation can be explained by a simple example in figure 2 which involves "an image (6x6x3), a kernel (3x3x3) where 3 indicates 3-channel in both aspects and bias (1x1)" [19]. Here, the operation is performed with zero padding and stride of one which means "the kernel is passed over every (3x3) segment of an image in a process of overlapping movement from left to right by moving a position for every operation. This results in a (4x4x3) convolved feature map" [19] and here, the number of input channels and kernel channels are always identical to each other.

Pooling Layers

CNN contains some pooling layers along with convolution layers. These layers can also be called subsampling layers. Here, outputs of the convolution layer act as input for pooling layers. Non-trainable kernels or windows are used in two-dimensional pooling layers to "down-sample input features" [19]. This process will reduce the size of the features significantly and helps in removing the network's position dependency. Types of subsampling are "average and max pooling" [19] which are widely used, and some others are "stochastic pooling, spectral pooling, spatial pyramid pooling, L2-norm pooling, and multi-scale orderless pooling" [11]. Generally, pooling layers consist of

pooling neurons that are ready to implement a large number of conventional operations simultaneously with average pooling and "L1/L2 normalization" as well [11].

This layer operates on "every activation map within the input and scales its dimensionality using a function-based" [11] on the opted pooling method, for max-pooling 'MAX' function is used and for average pooling 'AVG' function is used. Both the methods are computed "on maximum and average of the values present in each activation map or kernel respectively that should be provided for the feature map that comes as a result" [19]. The dimensions of this resulting feature map in the pooling layer are obtained as the same as the convolution layers. The following figure 3 shows how max and average pooling works in kernels and figure 4 shows a pooling/subsampling operation with a (6x6x3) image and a (2x2x3) kernel where the operation is done with an applied "stride of 2 and padding as 0 which means that the kernel is shifted for every (2x2) part of the input in a non-overlapping movement by moving 2 positions for every operation that gives a (3x3x3) feature map that is down-sampled or small-scaled." [19].

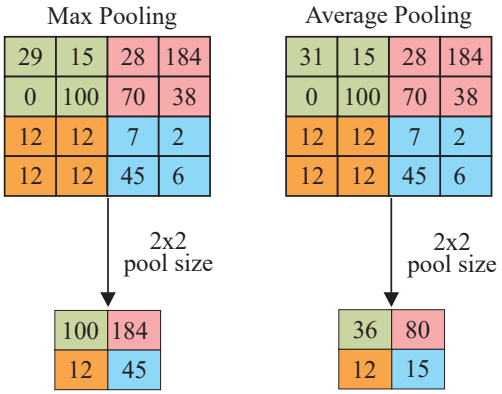


Figure 3. *Max pooling and Average pooling example*

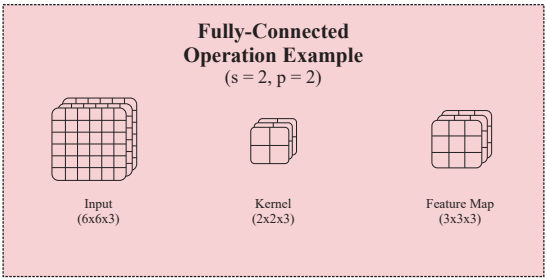


Figure 4. *Sub-sampling operation example*

Fully Connected Layer

Fully connected layers are so not unique to CNNs, and they are considered the last layers of most of the CNNs which appear after many convolutional and pooling operations are implemented. These layers are individual neural networks that contain single or many hidden layers, and perform operations are like "obtaining the product of their inputs and the trainable weight vectors like kernels or filters, and sometimes a trainable bias is added to the obtained results. The output obtained from these layers is passed through activation functions as same as that of convolution layers" [19].

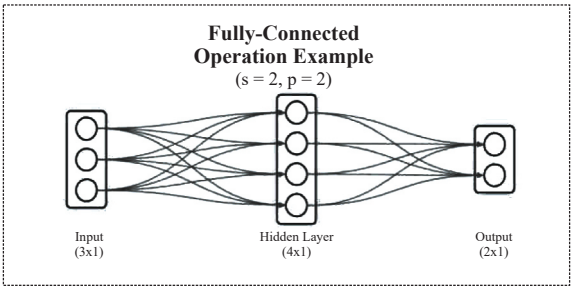


Figure 5. *Fully Connected Operation example*

A fully connected operation is shown in figure 5 that involves an input of dimensions (3x1), a hidden layer

of dimensions (4x1), and an output that is obtained with dimensions (2x1) [19]. In this layer, neurons have connections to all the activations present in the previous layer which can be observed in a multi-layer perceptron neural network and these activations will be computed with the help of matrix operation which can also be followed by a bias offset. "A fully connected layer provides a two-dimensional output to the output layer in which **softmax** function or **sigmoid** function can be used to predict the input class label" [11].

LeNet – 1

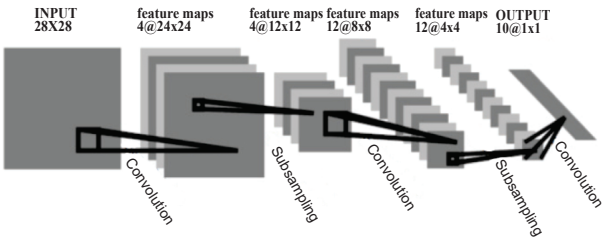


Figure 6. *LeNet-1 architecture*

LeNet-1 is a simple CNN that has five layers. This network is developed to consider "small, single-channel images of (28x28) size with an overall of 3,246 trainable parameters and 139,402 connections" [19]. This is originally trained on the USPS database as mentioned above in which the error rate is 1.7%. Throughout the testing process, the down-sampling of images is observed to a tensor of dimensions (16x16) and then centered in a larger tensor of (28x28) dimension. This whole process is completed with the help of a customized input layer [1] and the five layers of this LeNet-1 are as follows.

a. Convolutional layer

- Number of kernels = 4

- Kernel size = 5x5
- Padding = 0
- Stride = 1

The First Layer of LeNet-1 is a convolutional layer that considers a "(28x28) one channel image tensor as an input and zero padding is performed using (5x5) kernels with stride 1 as shown above which results in (24x24x4) output tensor" and that was "moved through **tanh** activation function since it can instigate non-linearity and then passed to next layer as input" [19].

b. Subsampling or Average Pooling Layer

- Kernel size = 2x2
- Padding = 0
- Stride = 2

This is the second layer with average function as pooling operation and considers output from the previous layer (24x24x4) tensor as input and operates zero-padding pooling operation using (2x2) kernel with stride 2 which results in (12x12x4) output tensor that acts as input to the succeeding layer.

c. Convolution Layer

- Number of kernels = 12
- Kernel size = 5x5
- Padding = 0
- Stride = 1

This is the third layer which takes (12x12x4) tensor as input and performs zero-padding with stride 1 considering 12 kernels of (5x5) and provides (8x8x12) output sensor

which acts as an input for the next layer after passing through a **tanh** activation function.

d. Subsampling or Average Pooling Layer

- Kernel size = 2x2
- Padding = 0
- Stride = 2

This is the fourth layer which takes (8x8x12) tensor, from the previous layer as input and performs zero-padding like the second layer using (2x2) kernel with stride 2 that results in (4x4x12) as output tensor that acts as input for the next layer.

e. Fully Connected Layer

- Output features = 10

This is the final layer that takes the flattened output (4x4x12) tensor of the previous layer and performs a "weighted sum operation with added bias term" [19] that generates a (10x1) output tensor which is sent through the **softmax** activation function gives output that has predictions of the network.

2. LeNet – 4

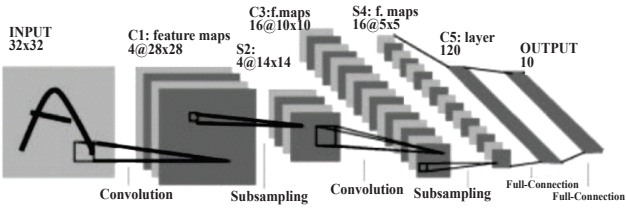


Figure 7. *LeNet-4 Architecture*

- LeNet-4 is a CNN with six layers and an improvement over LeNet-1. The main variation between LeNet-1 and LeNet-4 can be said as Lenet-4 is developed

to consider a larger image with a size of (32x32) whereas Lenet-1 is capable of considering an image with a size of (24x24). This network is based on "51,050 trainable parameters and 292,466 connections" [19]. Throughout the testing process, the images considered are down-sampled to (20x20) tensor and then a larger tensor of (32x32) is centered from that down-sampled. This network gives an "error rate of 1.1% on test data" from which it can be concluded that "there is a decrease of 35% when compared to Lenet-1" [19]. The six layers of this Lenet-4 are as follows:

a. Convolution Layer

- Number of kernels = 4
- Kernel size = 5x5
- Padding = 0
- Stride = 1

This is the first convolutional layer with an input tensor of (32x32x1) and performs zero-padding using kernels (5x5x4) with stride 1 which produces an output tensor (28x28x4) which is later moved through the **tanh** activation function and sent as an input for next layer.

b. Subsampling Layer

- Kernel size = 2x2
- Padding = 0
- Stride = 2

This is the second layer with average pooling and accepts output from the previous layer (28x28x4) tensor as an input and performs zero-padding using (2x2) kernel with stride 2 which results in (14x14x4) output tensor.

c. Convolution Layer

- Number of kernels = 16
- Kernel size = 5x5
- Padding = 0
- Stride = 1

This is the third layer which takes (14x14x4) tensor as input and performs zero-padding with stride 1 and using 16 kernels of (5x5) and provides (10x10x16) output sensor which acts as an input for the next layer after passing through the **tanh** activation function

d. Subsampling Layer

- Kernel size = 2x2
- Padding = 0
- Stride = 2

This is the fourth layer with average pooling and accepts output from the previous layer (10x10x16) tensor as an input and performs zero-padding using (2x2) kernel with stride 2 which results in (5x5x16) output tensor.

e. Fully Connected Layer

- Output features = 120

This is the fifth layer that takes the flattened output tensor of (5x5x16) of the previous layer and performs a "weighted sum operation with added bias term" [19] that produces a (120x1) output tensor which passes through the **tanh** activation function that acts as an input for the next fully connected layer.

f. Fully Connected Layer

- Output features = 10

This is the final layer that takes the output tensor of (120x1) of the previous layer as input and performs a weighted sum operation like the previous layer that produces a (10x1) output tensor which passes through the **softmax** activation function gives output that has predictions of the network same as that of LeNet-1

4. LeNet – 5

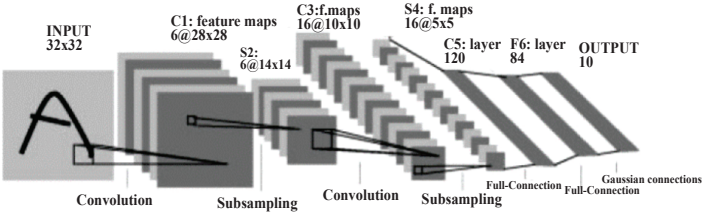


Figure 8. *LeNet–5 Architecture*

LeNet-5 is the improved LeNet-4 and the most popular of all three and this features several trainable layers. This network consists of an image sensor of (32x32) as input with a total trainable parameter of 60,850 and 340,918 connections. This network gets an "error rate of 0.95% on test data which explains that 44% decrease in error rate when compared to LeNet-1 and 13% decrease in error rate when compared to LeNet-4" [19]. LeNet-5 architecture has seven layers, and they are as follows:

a. Convolutional Layer

- Number of kernels = 6
- Kernel size = 5x5
- Padding = 0
- Stride = 1

This is the first convolutional layer with an input tensor of (32x32x1) and performs zero-padding using 4 kernels of (5x5) with stride 1 which produces an output tensor

(28x28x6) which is later passed through the **tanh** activation function and sent as an input for next layer.

b. Subsampling Layer

- Kernel size = 2x2
- Padding = 0
- Stride = 2

This is the second layer with average pooling and accepts output from the previous layer (28x28x6) tensor as an input and performs zero-padding using (2x2) kernel with stride 2 this subsampling operation is done by multiplying the average of the regions with trainable parameters and adding trainable bias to the results which produce in (14x14x6) output tensor, and this is passed through **sigmoid** activation function and then acts as an input for next layer.

c. Convolutional Layer

- Number of kernels = 16
- Kernel size = 5x5
- Padding = 0
- Stride = 1

This is the third convolutional layer with an input tensor of (14x14x6) and performs zero-padding using kernels of (5x5x16) rather than 6 with stride 1. This operation is performed on a section of the prior layer which produces an output tensor (10x10x6) which is later passed through the **tanh** activation function and sent as an input for the next layer.

d. Subsampling Layer

- Kernel size = 2x2
- Padding = 0

- Stride = 2

This is the fourth layer with average pooling and accepts output from the previous layer (10x10x16) tensor as an input and performs zero-padding using (2x2) kernel with stride 2 and this subsampling operation is done which produces in (5x5x16) output tensor and this is passed through **sigmoid** activation function and then acts as an input for next layer.

e. Fully Connected Layer

- 7. Output features = 140

This is the fifth layer that takes the flattened output tensor of (5x5x16) of the previous layer and performs a "weighted sum operation with added bias term" [19] that produces a (120x1) output tensor which passes through the **tanh** activation function that acts as an input for the next fully connected layer.

f. Fully Connected Layer

- Output features = 140

This is the sixth layer that takes the output tensor of (120x1) of the previous layer and performs the same operation as the previous layer that produces an (84x1) output tensor that passes through the **tanh** activation function that acts as an input for the next fully connected layer.

g. Fully Connected Layer

- Output features = 140

This is the final layer that takes the output tensor of (84x1) of the previous layer as input and performs a weighted sum operation like both previous layers that produce a (10x1) output tensor which passes through the **softmax** activation function gives output that has predictions of the network same as that of LeNet-1 and Lenet-4.

III. Applications of LeNet

Lung Cancer Detection

For Lung cancer detection a hybrid model was proposed in which "LeNet, AlexNet, and VGG-16" [2] are used. In LeNet architecture, there is a "(5x5) filter in which the image size varies from (32x32x1) to (28x28x6). In addition to this operation, the features obtained from the last fully connected layer are then applied with different models like Linear Regression (LR), Linear Discriminate Analysis (LDA), Decision Tree (DT), Support Vector Machine (SVM), k-nearest neighbor (kNN) and then softmax" [2]. All these classifiers are tested at the end of the task performance and examined individually by comparing their performances. For increasing the classification accuracy, the image augmentation technique is also used in the process of training these models. By doing image augmentation, "20 additional images were gathered from each original image in the dataset". For LeNet architecture, "ADAM and RMSProp optimization methods" [2] are used while developing the model. To examine the outcome, "sensitivity, specificity, accuracy, receiver operator characteristics curve (ROC), area under the curve (AUC)" are used [2]. In lung nodule classification, LeNet is used along with AlexNet and VGG-16, and this model is evaluated using specificity, sensitivity, and accuracy. The values obtained by using these combined models are" accuracy of 99.51%, the sensitivity of 99.32%, and specificity of 99.71%". [7]

Brain Tumor Detection

Brain tumor detection is done with the help of "Magnetic Resonance Imaging (MRI) and this is commonly used in medical imaging and image processing" in the process of diagnosing the dissimilarity in different parts of the body. By using LeNet-5, detection of brain cancer can be done

with 99% accuracy, and it is also used in the "classification of Alzheimer's disease" in the brain. For this Alzheimer's disease classification, the features used are shape and scale and this LeNet architecture achieved the "possibility to generalize a method to predict different stages of Alzheimer's for different age groups" [3] and this gives an accuracy of 96.85% [3].

Identification of Corona Virus

The novel coronavirus is declared a pandemic which is started in China. "For the detection of coronavirus and SARS_MERS infected patients" [4] and for distinguishing the patients from healthy subjects, CNN based models are used including LeNet-5 and for this "lung X-ray scans" are used which is proven to be the most challenging task because of the overlap in characteristics of different coronavirus types [4].

Classes	Number of images
nCov	45
SARS_MERS	49
Normal	51
Total	145

Table 1. *Description of image dataset used.*

Evaluation metrics that are used for the "X-ray image classification" are precision, recall, F1-score, and accuracy. These values can be calculated from the following expressions

$$Precision = \frac{True\ Positive\ (T_p)}{True\ Positive\ (T_p) + False\ Positive(F_p)}$$

$$Recall = \frac{True\ Positive\ (T_p)}{True\ Positive\ (T_p) + False\ Negative(F_n)}$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}$$

By using the above expressions, an evaluation of the model is done, and the evaluation metrics of the developed model are as table 2 [4].

	Precision	**Recall**	**F1 score**
nCov	100	70	82
Normal	86	100	92
SARS_MERS	81	93	87
Weighted mean accuracy	88	87	86

Table 2. Numerical results of LeNet-5

Diabetic Retinopathy (DR) Detection

"Diabetic Retinopathy" is a complication observed commonly in diabetes mellitus which is a reason the cause of "lesions on the retina" which affects vision. If this lesion is not detected, there is a problem that leads to blindness and this "DR is not a reversible process" [5]. This DR detection is done based on DR retina fundus images. In this classification process, evaluation metrics considered are specificity, sensitivity, accuracy, and AUC. The expressions are as follows

$$Specificity = \frac{True\ Negative\ (T_n)}{True\ Negative\ (T_n) + False\ Positive(F_p)}$$

$$Recall = Sensitivity = \frac{True\ Positive\ (T_p)}{True\ Positive\ (T_p) + False\ Negative(F_n)}$$

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}$$

By using the above expressions, evaluation of the model is done, and the evaluation metrics of the developed model are as table 3. "Y. Yan et al. detected DR red lesions in DIARETB1 dataset" [5] using LeNet architecture with Random Forest classifier after integrating the improved pretrained LeNet and handcrafted features. The noise in the images is filtered using a "Gaussian filter and a morphological method" is used. After this process, U-net CNN architecture is used for "blood vessel segmentation" from the images. This improves LeNet has "4 convolutional layers, 3 max-pooling or subsampling layers and only one fully connected layer" [5].

	AUC	**Accuracy**	**Sensitivity**	**Specificity**
CNN (Improved LeNet, U-net)	CPM = 0.4823	-	48.71%	-

Table 3. The method used for DR detection with evaluation metrics

For lesion detection, CNN can be "trained using LeNet architecture having an introduction of a sliding image while testing so that multi-class outcome probability can be obtained" [10].

Medical Image Processing

For "medical image segmentation and classification, deep learning architectures" are widely used. In this medical image processing, detection of Diabetic retinopathy, Alzheimer's, histological and microscopical elements can also be done through which malaria can also be detected. Dong and Bryan [6] are the authors who "developed four systems to detect infected and non-infected cells by malaria using CNN models along with SVM". The CNN

architectures that are used here are "GoogleNet, LeNet-5, and AlexNet which are used for automatic feature extraction and also classification" [9].

By using these architectures, the accuracy obtained is 98.13%, 96.18%, and 95.79% respectively for the three architectures mentioned above. By using SVM it achieved the lowest accuracy of 91.66%. For Tumor detection, Gastrointestinal disease also, these architectures are used. Sarraf et al. [6] mentioned that "scale and shift of the features like shape of the data, mean and standard deviation of the data can be carried out for classifying functional MRI 4D". Here, the system is "trained on 270,900 images and validated and tested on 90,300 functional MRI images" [9]. By developing this system, the author obtained 96.86% of accuracy in identifying Alzheimer's.

Medical Image Analysis

In "medical image analysis, Deep Convolution Neural Network (DCNN) is usually used to learn features from the underlying data" [8]. For this, a CNN architecture like LeNet-5 is used for the "classification of medical images with N classes having a patch of (32x32) from original 2D image". This can be observed in the figure below as it contains "convolutional layers, max-pooling layers, and fully connected layers". Here, "each convolutional layer generated different feature maps with different sizes and pooling layers reduce the size of the maps that are to be transferred to the next layers" [8]. "The fully connected layers have the output of the predicted class". Some of the recent studies clearly show that deep learning algorithms can be widely used for "medical image segmentation, computer-aided diagnosis (CAD), detection of diseases, classifications and medical image retrievals" as well [8].

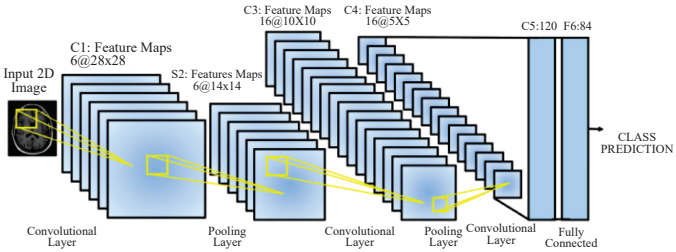


Figure 9. *LeNet architecture for medical image classification.*

Sleep Apnea Detection

Sleep apnea is a commonly seen respiratory sleep disorder that can lead to serious neurological and cardiovascular diseases. For detecting this disease, the "ECG signal is one of the most relevant signals of this disease's occurrence". For this detection, two separate datasets are considered of which one is the "physionet apnea ECG dataset which contains a total of 70 single ECG signals and the other is the UCD dataset which is from University College Dublin" [12]. The figure below depicts the architecture of the SA detection using LeNet-5.

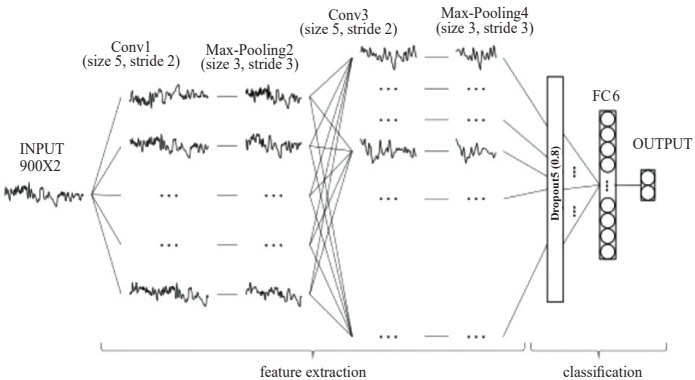


Figure 10. *The architecture of modified LeNet-5*

The details of the modified LeNet-5 CNN can be explained as follows.

Layer	Parameter	Output Shape	Number
Input	-	(None, 900, 2)	0
Convl	32 x 5 x 2, stride 2, pad 0	(None, 448, 32)	352
Max pooling2	3, stride 3, pad 0	(None, 149, 32)	0
Conv3	64 x 5 x 2, stride 2, pad 0	(None, 73, 64)	10304
Max polling4	3, stride 3, pad 0	(None, 24, 64)	0
Dropout5	0.8 rate	(None, 24, 64)	0
FC6	32, relu	(None, 32)	49184
Output	2, softmax	(None, 2)	66

Figure 11. *Layers of modified LeNet-5 architecture*

From the above, it can be understood how the "convolutional layer, max-pooling layers, and fully connected layers" are considered along with stride. For evaluation of this model, specificity, sensitivity, and accuracy are considered. The overall performance of the model is shown in figure [12].

Layer	Accuracy(%)	Sensitivity(%)	Specificity(%)	AUC
SVM	81.4	76.9	84.3	0.887
LR	80.8	75.7	84.0	0.884
KNN	77.5	68.1	83.4	0.826
MLP	81.1	71.3	87.2	0.898
LeNet-5	87.6	83.1	90.3	0.950

Figure 12. *Evaluation metrics for modified LeNet-5*

Conclusion

In this paper, the architecture of LeNet is discussed concerning its previous versions and explained some of its applications in medical sciences. LeNet is the first created CNN and is best to use for computer vision and Alzheimer's detection along with sleep apnea and lung cancer problems. LeNet-5 is used in combination with Alexnet, GoogleNet and also along with AlexNet, VGG-16.

References

[1] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition,"

in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998

[2] doi: 10.1109/5.726791.

[3] Sathyakumar K, Munoz M, Singh J, et al. (August 25, 2020) Automated Lung Cancer Detection Using Artificial Intelligence (AI) Deep Convolutional Neural Networks: A Narrative Literature Review. Cureus 12(8): e10017. doi:10.7759/cureus.10017

[4] Arabahmadi, M.; Farahbakhsh, R.; Rezazadeh, J. Deep Learning for Smart Healthcare—A Survey on Brain Tumor Detection from Medical Imaging. Sensors 2022, 22, 1960.

[5] https://doi.org/10.3390/s22051960

[6] Mohammed Chachan Younis, Evaluation of deep learning approaches for identification of different corona-virus species and time series prediction, Computerized Medical Imaging and Graphics, Volume 90, 2021, 101921, ISSN 0895-6111, https://doi.org/10.1016/j.compmedimag.2021.101921.

[7] Wejdan L. Alyoubi, Wafaa M. Shalash, Maysoon F. Abulkhair, Diabetic retinopathy detection through deep learning techniques: A review, Informatics in Medicine Unlocked, Volume 20, 2020, 100377, ISSN 2352-9148

[8] https://doi.org/10.1016/j.imu.2020.100377.

[9] Razzak M.I., Naz S., Zaib A. (2018) Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In: Dey N., Ashour A., Borra S. (eds) Classification in BioApps. Lecture Notes in Computational Vision and Biomechanics, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-319-65981-7_12

[10] Thakur, S.K., Singh, D.P. & Choudhary, J. Lung cancer identification: a review on detection and classification. Cancer Metastasis Rev 39, 989–998 (2020).

[11] https://doi.org/10.1007/s10555-020-09901-x

[12] Anwar, S.M., Majid, M., Qayyum, A. et al. Medical Image Analysis using Convolutional Neural Networks: A Review. J Med Syst 42, 226 (2018). https://doi.org/10.1007/s10916-018-1088-1

[13] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, Volume 42, 2017, Pages 60-88, ISSN 1361-8415, https://doi.org/10.1016/j.media.2017.07.005.

[14] Sourya Sengupta, Amitojdeep Singh, Henry A. Leopold, Tanmay Gulati, Vasudevan Lakshminarayanan, Ophthalmic diagnosis using deep learning with fundus images – A critical review, Artificial Intelligence in Medicine, Volume 102, 2020, 101758, ISSN 0933-3657, https://doi.org/10.1016/j.artmed.2019.101758.

[15] Sakib, S.; Ahmed, N.; Kabir, A.J.; Ahmed, H. An Overview of Convolutional Neural Network: Its Architecture and Applications. Preprints 2018, 2018110546 (doi: 10.20944/preprints201811.0546.v4).

[16] Wang T, Lu C, Shen G, Hong F. 2019. Sleep apnea detection from a single-lead ECG signal with automatic feature extraction through a modified LeNet-5 convolutional neural network. PeerJ 7:e7731 https://doi.org/10.7717/peerj.7731

[17] https://pyimagesearch.com/2016/08/01/lenet-convolutional-neural-network-in-python /# :~:text=The%20LeNet%20architecture% 20was%20first%20introduced%20by%20LeCun,primarily%20for%20OCR%20and%20character%20recognition%20in%20documents.

[18] https://towardsdatascience.com/understanding-and-implementing-lenet-5-cnn-architecture-deep-learning-a2d531ebc342

[19] https://www.kaggle.com/code/blurredmachine/lenet-architecture-a-complete-guide/notebook

[20] http://d2l.ai/chapter_convolutional-neural-networks/lenet.html

[21] https://en.wikipedia.org/wiki/LeNet#:~:text=From%20Wikipedia%2C%20the%20free%20encyclopedia%20LeNet%20is%20a,LeNet-5%20and%20is%20a%20simple%20convolutional%20neural%20network.

[22] https://sh-tsang.medium.com/paper-brief-review-of-lenet-1-lenet-4-lenet-5-boosted-lenet-4-image-classification-1f5f809dbf17

[23] https://towardsdatascience.com/understanding-lenet-a-detailed-walkthrough-17833d4bd155

A Novel Approach of Diabetes Treatment Management System Using Artificial Intelligence

Narra Harshini

Ug scholar/ IT
Alliance College of Engineering and Design
Bangalore,india.
narraharshini15@gmail.com

Rahul Ray

Ug scholar/ IT
Alliance College of Engineering and Design
Bangalore,India.
jrahulray7@gmail.com

Dr. Senbagavalli .M

Associate professor/ IT
Alliance College of Engineering and Design
Bangalore,india.
senbagavalli.m@alliance.edu.in

***Abstract*—** Humanity is still suffering from deadly diseases and some of these diseases are still not curable.

The solution for this is to continue our research in methods of preventing these diseases. In this article, we can see the development taking place in AI on the deadly disease Diabetes. This article has information about the on-going research about medical science using AI which supports the fight against different types of deadly disease diabetes. In this AI system and research, we are using statistics and machine learning to find methods for fighting against this disease. Using statistics, we even derived the numbers of articles released in recent years. With respect to the current situation in the global world, we can expect and predict the main role AI is going to play and how the data will be organized to predict a certain performance.

Keywords— *Artificial intelligence, Machine learning, Statistical Analysis and Deep Learning*

I. INTRODUCTION

How is artificial intelligence useful in therapeutic sciences? It is artificial intelligence that points to precisely made progressed induction on an expansive sum of data. The huge companies like Microsoft, Apple, Google and IBM are working on making objects utilizing artificial intelligence in the healthcare field, and also, they contribute vast sums for development in the field of artificial intelligence to make personalized search engines for virtual personal assistants. Artificial intelligence-based investigation could be a substitution of ordinary statistical strategies, i.e., Linear regression and logistic regression. In pharmaceuticals, both deduction and forecast are critical objectives that both routine measurements and machine learning can accomplish. Insights emphasizes giving a system for decision-making through induction. Artificial intelligence moreover makes a difference to prepare all the data for human wellbeing. On the other hand, machine learning is utilized to maximize the

execution of anticipating answers to questions adversary which we do not know yet. In any case, it is known that manufactured insights may or may not be great sufficient to deliver 100% exact counsel to specialists, almost the finest strategy of treatment.

II. DIABETES

Sometimes recently looking into what diabetes is, let us see why it is imperative to be concerned about this disease. Diabetes depicts disease of anomalous carbohydrate digestion systems that are characterized by hyper-glycemia. It is related with relative or supreme disability in affront discharge, alongside shifting degrees of fringe resistance to the movement of affront. It affects the retina of the eye first and sometimes does blood clotting in the eye and makes a blurred view. Each time, diabetes is re-evaluating the present proposals for classifying, concluding and to screen diabetes, verifying unused data from inquiring about it and clinical practice.

So, it is clear that diabetes needs the same work in the field of research for creating artificial intelligence in finding arrangements how to treat this infection. Keeping up a low blood glucose level, one can prevent these hazards and its complications or in other words it ensures a long and satisfying life. Presently since we are stuck on the concept of diabetes, we should be able to examine how the advancement and investigation should take place within the field of AI. So, these will assist us in finding arrangements to the issues.

III. LITERATURE SURVEY

[1] the response of a patient can be evaluated by measuring the number of antibodies to various cell antigens in a body. By considering the present diabetes disease,

patients can be identified through immune related cases. [2] the rapid growth of diabetes is in the peak range. especially to the people who are close to their puberty age. To predict the difference between these types of diabetes, their genetic and clinical features are being examined. [3] here, in this they have introduced a diagnosis system. This system combines the new feature selection which helps to classify and detect type 2 diabetes by processing the image and text. [4] to predict the present diseases i.e., diabetes we need an efficient model which represents the human condition regarding the glucose intake. But these models do not give 100% efficient and required results. So, regarding this issue, we can introduce mobile platforms like apps to measure glucose level in a patient. [5] from the past many years, people worldwide are trying to introduce and develop a few systems which can control the blood sugar levels in a patient with the types of diabetes. This technology is mentioned as artificial pancreas.

[6] States that due to the rapid increase of the disease, there were chatbots developed with the help of artificial intelligence. There are many healthcare applications that are providing patients with some help regarding the disease. These AI chatbots were enabling more interactions with patients. [7] Usually we test the data under supervision of a specialist. We use this data to understand natural language like machine learning which helps in identification of disease. There was a prototype version as MIRA, which checks the performance in terms of accuracy, sensitivity etc.

[8] There was a mass usage of conventional agents which worked based on the input of natural language for health purposes. [9] Under machine learning, deep learning techniques was used to get the best results possible by image recognition and with the help of natural language processing. It was widely used in medicine and cybersecurity. [10]

There were several processes which were introduced to support care for Diabetes malady. Such processes include the concept of the artificial neural network, concept of the naïve bayes, concept of decision tree and k-nearest algorithm. ML has played a major role in creating screening for blood glucose variability.

IV. AI FOR DIABETES

These days there's no culminating AI framework that can confront and make sense of the existing issue of diabetes. Diab next company presented instruments based on insight calculations which are ready to offer help in assistance of diabetes treatment. In this process, patients can communicate through the Diab next framework which permits them to use and gadget which is interconnected to the web. This results in voice-activated communication. Especially doctors can visualize affront infusion measurement, pharmaceutical information, blood glucose levels, understanding nourishment and calculate carbohydrate tallies, the test comes about patterns in hemoglobin, diabetes sequencing profiles, help in additionality of changes made in work out and weight information. The framework is capable of foreseeing and controlling the blood sugar levels in individuals with information regarding diabetes with the help of biometric sensors.

V. STATISTICS AND MACHINE LEARNING

In statistics, we can estimate population values using data based on assumptions and then we can perform hypothesis tests to assess the difference between groups. Machine learning usually contains labelled data. And from this labelled data, we can judge an algorithm even if the input data is large. It can even optimize the input data to give out better and significant output. And in case if any new data is produced, then it will label it as much as possible. In fact,

we can conclude that machine learning and statistics are the same because both of these concepts produce output or results based on the available data. And when talking about the disease, we could say that preventing a problem is better than fighting against it. [Fig.1] As machine learning uses labelled data to give out best output possible, it is being a field with greater development. So, by using the concepts of machine learning and statistics, we can predict which disorder or problem a person is going through. Using the data, it can even predict if a person is going through diabetes disorder or not. According to recent research, scientists have concluded that 70% of the patients are suffering from type-2 diabetes.

This type-2 diabetes is due to diet-related and genes play a major role in likelihood. The main thing to solve these problems regarding diabetes is we should have a fit model. We can say machine learning plays a major role in this, only because of its unique role it plays in the present generation. Its challenging character makes it more unique in the medical science field. Considering the data present, we can judge which concept is better to get certain output to predict which medicine. And hence there is much research going on to find the right medicine and cure.

VI. AI USE IN CURRENT DIABETES MANAGEMENT AND FUTURE WORK

Mainly, AI is used to spot the patterns of behaviour which lead us to the information that the sugar levels are high or low in patients. In-this process, the glucose monitors used by patients collect the data and then work efficiently. And as a well-known method, the decision tree is one of the most popular techniques used in the field of medical science. And presently there are many food administrations which are using AI based devices to cure diseases efficiently.

Instead, overfitting is the main issue occurring because the predicted accuracy is high. Even though machine learning gives perfect output or results from the provided labelled data, it also has some problems and disadvantages in this field.

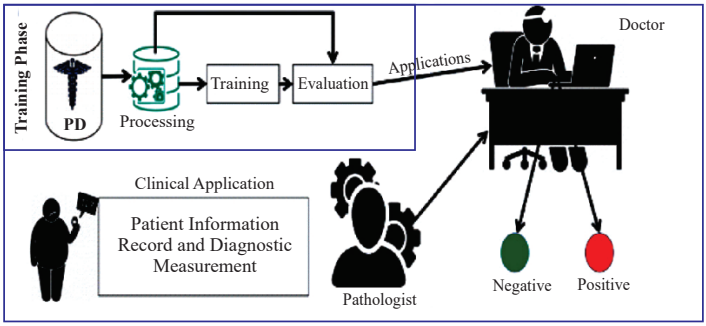


Fig.1: *machine learning in diabetes [12]*

It is not able to predict the new diagnosis by monitoring. Still, it is very useful for predicting performance when compared to statistics. It can tell that taking proper decisions about diet can prevent from type 2 diabetes. And certain apps can be referred to access the health condition of the patient. And we can use apps for calories measuring as it is one of the main-pictures acting in the deadly disease diabetes. And evaluating patients time-to-time about the condition can also be a great check. Remote monitoring can decrease the amount of time used to analyse the patient. And improving data sets can give best results for the prediction and prevention of diabetes.

Further work is to study and develop systems using neural networking in deep learning.

VII. CONCLUSION

The above article is written based on the previously published one's on the development of AI to solve deadly

disease diabetes. Statistics of previously published articles, number of people being affected due to diabetes and the efficient modes of treatments are being mentioned above in this article. As mentioned above, use of the latest technology and the development in the technology to fight diabetes is being one of the main reasons for its uniqueness. Following up certain latest technology like apps etc., to keep diabetes in balance. And trying to maintain the correct data set will help in predicting the possible efficient results, which might help for the prevention of diabetes.

REFERENCES

[1] Merger SR, Leslie RD, Boehm BO. The broad clinical phenotype of Type 1 diabetes at presentation. Diabetes Med 2013; 30:170.

[2] Karjalainen J, salmela P, Ilonen J, et al. A comparison of childhood and adult type I diabetes mellitus. N Engl J Med 1989; 320:881.

[3] Bhuvaneswari, G., & Manikandan, G. (2018). A novel machine learning framework for diagnosing the type 2 diabetics using temporal fuzzy ant miner decision tree classifier with temporal weighted genetic algorithm. Computing.

[4] Reymann, M. P., Dorschky, E., Groh, B. H., Martindale, C., Blank, P., & Eskofier, B. M. (2016). Blood glucose level prediction based on support vector regression using mobile platforms. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS.

[5] Bertachi, A., Ramkissoon, C. M., Bondia, J., & Vehí, J. (2018). Automated blood glucose control in type 1 diabetes: A review of progress and challenges. Endocrinologia, Diabetes y Nutricion

[6] Schachner, T., Keller, R., and V Wangenheim, F. 2020. "Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review," Journal of Medical Internet Research (22:9), p. e20701.

[7] Rehman, U. U., Chang, D. J., Jung, Y., Akhtar, U., Razzaq, M. A., and Lee, S. 2020. "Medical Instructed Real-Time Assistant for Patient with Glaucoma and Diabetic Conditions," NATO Advanced Science Institutes Series E: Applied Sciences (10:7), Multidisciplinary Digital Publishing Institute, p. 2216

[8] Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., and Coiera, E. 2018. "Conversational Agents in Healthcare: A Systematic Review," Journal of the American Medical Informatics Association: JAMIA (25:9), pp. 1248–1258

[9] Grzybowski, A., Brona, P., Lim, G., Ruamviboonsuk, P., Tan, G. S. W., Abramoff, M., and Ting, D. S. W. 2020. "Artificial Intelligence for Diabetic Retinopathy Screening: A Review," Eye (34:3), pp. 451–460.

[10] Ellahham, S. 2020. "Artificial Intelligence: The Future for Diabetes Care," The American Journal of Medicine (133:8), pp. 895–900.

[11] FIG.1-Source: https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.thelancet.com%2Fjournals%2Flandia%2Farticle%2FPIIS2213-8587(21)00208-4%2Ffulltext&psig=AOvVaw0Od7qGjQlhrYYQeH0Su7ps&ust=1647762620936000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCIDQj6LY0fYCFQAAAAAdAAAAABAD

[12] FIG.2 – Source:

[13] https://www.google.com/url?sa=i&url=https%3A%2F%2Fgithub.com%2Fkamruleee51%2FDiabetes-Prediction-Using-ML-Classifier

s&psig=AOvVaw1rthkOK3ePwDVql3bpAJmb&ust=1647956691728000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCPD6v56r1_YCFQAAAAAdAAAAABBW

Diet and Exercise Tracking for Daily Requirements Using Artificial Intelligence Technique

Vasavi Annapureddy[1], Keerthana J[1*], Mohit Hemanth Kumar[2]

[1]Department of Computer Science and Engineering
[2]Department of Mechanical Engineering
Alliance College of Engineering and Design, Alliance University, Bengaluru, Karnataka, India.
*Corresponding Author Email: jkeerthana2004@gmail.com

***Abstract*:** Recently, there has been a concern about our looks and fitness freaks in today's world. The consumption of more food leads to gaining weight in the long run if it exceeds the daily requirement. The present investigation presents the control of calories consumption using the artificial intelligence (AI) model. Based on the AI model, the correct quantity of food is recommended and suggests the necessary workouts to reduce the extra calories. From this investigation, it can be concluded that people can still eat any food and do essential exercises needed for daily requirements to maintain proper health.

Keywords: Artificial intelligence; Calories; Diet plan; Exercise tracking; Workout tracking

INTRODUCTION

The living standards of people are increasing day by day and that leads to increasing demand for health in one's life. But due to lack of time, no one can give their time for their health-related issues. There is an increasing demand for maintaining a perfect body in everyone. Inspired by this we have developed the idea of this software [1-2]. This software scans the image of the food item which we are eating and gives insight about what is the number of calories which is present it in. People tend to have desires about the food they want to eat. but tend to cut down because it would cause gain weight. but this project enables the user to have whatever they want to without getting much concerned about weight gain. this project uses AI to keep track of your diet plan and gives insights about what to eat and how much to eat [3-4].

Heading towards the modern era, Artificial Intelligence (AI) can address the difficulties of this new worldview. The major goal of this project is to use artificial intelligence to get an insight into calories to maintain a person's healthy lifestyle. This project takes an idea about the age group of the person so that it can manipulate the data as per the age group. It also takes all his/her physical data such as weight height and whether they are currently having health problems such as B.P., thyroid, diabetes, etc. It can give suggestions to them on the required calorie intake for the day. And also, the required food items they can cook for themselves to meet their daily requirement. Whenever the user uploads the image, it detects the number of calories and adds it to the calorie intake of the person. If the calorie intake exceeds the required calorie intake, it suggests the

required exercises (demo videos) burn the extra calories. This software also keeps track of water intake which the user requires to drink per day and the number of hours one needs to sleep [5]

DATA INPUT

The software first takes input from the user for inputting their age so that it will display the content and manipulate it according to the age group of the user. Because the amount of calorie intake is different for the different age groups. The required calorie intake for each person is stored in the database and checked with it.

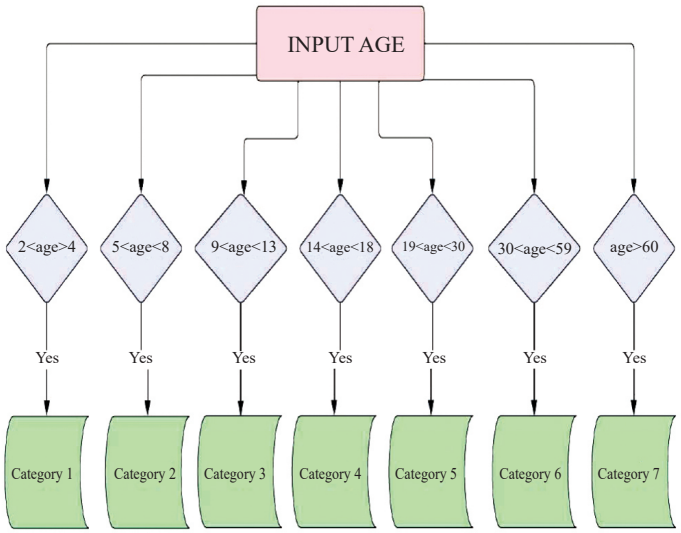


Figure 1*: category distribution according to age*

The user then enters all the physical parameters such as gender, height, weight, and also information about the health issue they are facing such as obesity, blood pressure, diabetes, thyroid, and many more such problem so that we can alter their diet plan accordingly. This will also enable

the user to take care of their diet by the disease they are having. For example, the software will suggest the required amount of sugar intake for the person affected with diabetes.

Table 1 Details of calories required according to different age categories

CATEGORY	CALORIE INTAKE (MALE)	CALORIE INTAKE (FEMALE)
2-4	1000-1600	1000-1400
5-8	1200-2000	1200-1800
9-13	1600-2600	1400-2400
14-18	2000-2600	1800-2200
19-30	2400-3000	2000-2400
31-59	2200-3000	1800-2200
60 ABOVE	2000-2600	1600-2000

The data fed by the user is compared with the following data table and the required amount of calorie requirement is told to the user. Next, if the user chooses to display some of the food items according to their daily requirement the necessary food items are displayed with which the user can plan out their meals for the day.

IMAGE PROCESSING USING ARTIFICIAL INTELLIGENCE (AI)

OpenCV

It is an open-source built-in library used for various computer vision and machine learning applications. This library has more than 2500 algorithms of machine learning and computer vision. These algorithms can be used to detect faces, recognize faces, classify human actions in videos, identify objects, objects moving, extract

3D models of objects, find similar images from an image database, track camera movements, remove red eyes from images taken using flash, stitch images together to produce a high-resolution image of an entire scene, follow eye movements, recognize scenery, etc. It is available in languages like Python, C++, MATLAB, and Java interfaces and supports Windows, Linux, and Mac OS.

Database

The information regarding the features of the food item will be stored in a backend database (say, SQL). The features retrieved will be compared with this data and give the number of calories it has.

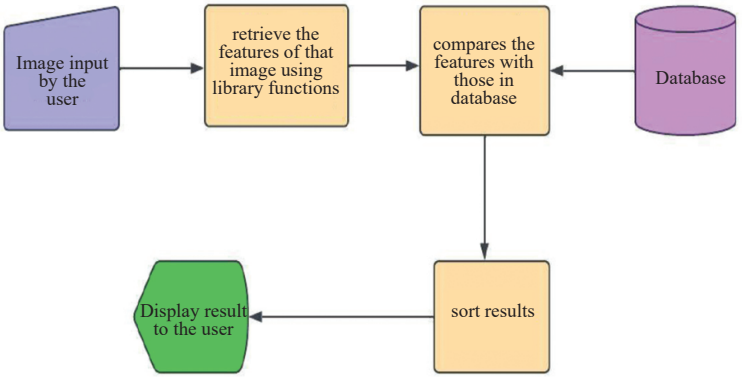


Figure 2*: Working of OpenCV*

Total calorie of food product=Calorie in the database * the quantity taken

Table 2. Details of calories of different kind of food items

FOOD ITEM	CALORIES
Idli (1 piece)	39

Vada	73
Dosa	168
Uttapam	92
Poori	101
Ragi mudde	107
Ragi roti	116
Rice	130
Dal	198
Rasam	59
Curd	98
Pickle	11
Ladoo	185
Chutney	45
Sambhar	2000
Upma	209
Pongal	212
Bread and Jam	112
Bread and Butter	67
Roti	71
Egg	155
Golgappa	36
Pizza	107
Pav bhaji	400
Pakoda	315

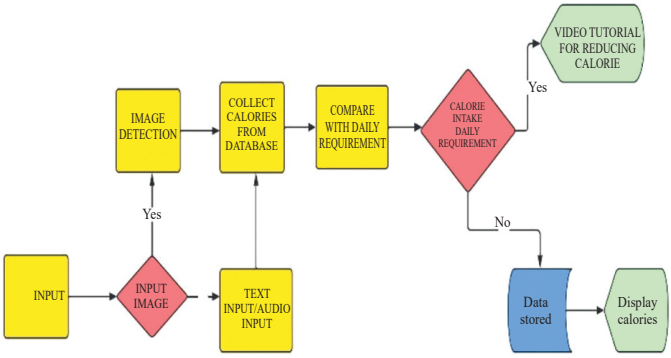


Figure 3*: overall working of the software*

Workout

Based on the data fetched by the software if your calorie intake is greater than the daily requirement the software would suggest you do some exercises with a video tutorial (from YouTube) by searching the keyword fetched from the database, based on which workout will reduce how many calories.

Table 3. Details of calories during the physical exercise

EXERCISE	CALORIES BURNT
Walking (1 km)	25-30
Running (1 km)	62
Jumping (100)	14
Sit-ups (100)	15
Cycling (1 hr)	450-750
Rock climbing (1 hr)	500-900
Push-ups (1 minute)	7

Water Tracking

Water is essential for the well-being of a human being. The human consists of 72% of water. Therefore, it is very essential to keep a track of water consumption. This software will remind user about drinking water at a particular time and keeps track of how much intake is done. An average human requires 3.7L of water per day.

Water required= persons weight * 0.033

Table 4. Amount of water to be drunk according to the age

Age	Water
9-13	7-8 cups
14-18	8-11 cups
Men 19+	13 cups
Women 19+	9 cups

Sleep Tracking

Sleep affects a person's health not only physically but also mentally, the approximate time an average person should sleep is 7 hours. The number of calories burnt in 7 hours of sleep is 420 calories. This software also suggests the best time for the user to go to sleep according to their age patterns. The software notifies the user to go to sleep at the required time.

Table 5. Amount of sleep to be taken based on the category of age

Age	Sleep (in hours)
3-5	11-13
6-13	9-11

14-17	8-10
18+	7-9

CONCLUSION

This software enables the user to take their food without worrying about what it would cost to their health. Not only that it also takes care of various factors contributing to a healthy lifestyle of a person which includes good sleep and proper hydration. As technological advancements take place in the field of technology so will be the lifestyle of an individual, artificial intelligence (AI) is contributing widely over many things because people prefer smart and fast things. To let them not forget their health in their ongoing busy lifestyle this software will guide them.

REFERENCES

[1] Image Processing Based Classification of Energy Sources in Eatables Using Artificial Intelligence Manjunathan A1*, Lakshmi A2, Ananthi S3, Ramachandran A4, Bhuvaneshwari C5 R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 7401 - 7407 Received 16 February 2021; Accepted 08 March 2021.

[2] Food Adulteration Detection using Artificial Intelligence: A Systematic Review Kashish Goyal1 · Parteek Kumar1 · Karun Verma1 Received: 15 September 2020 / Accepted: 27 April 2021 / Published online: 15 June 2021 © CIMNE, Barcelona, Spain 2021

[3] Development of a Mobile Application Platform for Self-Management of Obesity Using Artificial Intelligence Techniques Sylvester M. Sefa-Yeboah, 1 Kwabena Osei Annor, 1 Valencia J. Koomson,2 Firibu K. Scalia,3 Matilda Steiner-Asiedu,3 and Godfrey A. Mills 1 International Journal of Telemedicine and Ap-

plications Volume 2021, Article ID 6624057, 16 pages https://doi.org/10.1155/2021/6624057

[4] Modelling Techniques to Improve the Quality of Food Using Artificial Intelligence Varsha Sahni, 1 Sandeep Srivastava, 2 and Rizwan Khan 3 Hindawi Journal of Food Quality Volume 2021, Article ID 2140010, 10 pages https://doi.org/10.1155/2021/2140010

[5] A critical review on computer vision and artificial intelligence in food industry Vijay Kakani a, Van Huan Nguyen b, Basivi Praveen Kumar c, Hakil Kim a,*, Visweswara Rao Pasupuleti d,e,** journal homepage: www.journals.elsevier.com/journal-of-agriculture-and-food-research/

Review on the Role of Deep Networks in Autism Related Tasks

Rolla Subrahmayam

Ph.D-ongoing, UOH,

Jahanara Shaik

Ph.D-ongoing, Alliance University

***Abstract*—**The goal of artificial intelligence, specifically ma- chine learning, is to reduce the time and effort of humans by understanding and to learn intellectual and complex tasks using the data available. AI systems have successfully performed many tasks that need expertise at a large scale. The branch of artificial neural networks and deep learning is at the predominant and progressing phase, in many sectors, that have never been seen. The health sector is one domain that gets benefitted at a large scale. Many recent deep architectures have made analyses, diagnoses, classification, and predicting diseases and provided us at least human-level performance. In this paper, we provided a detailed review and comment on the neural and deep neural models in the health sector by considering the complex neurodevelopmental disorder of autism**.**

***Keywords*—**Autism, Neural Networks, CNN, loss, detection.

INTRODUCTION

ARTIFICIAL Intelligence is a branch of science that, in general, aims in mimicking human behavior. There are several definitions for artificial intelligence and all are agreeable but do not have a universally accepted definition [24]. Creativity and intelligence are the invaluable assets that humans possess and hence the human species dominated the earth compared to all others. So, there are consistent efforts to automate or create intelligence using machines for ages. Al- though there are several setbacks for artificial intelligence such as Godel's incompleteness theorem [17], progress in the task is phenomenal in modern times [4][7] [13][51]. Machine learning, reinforcement learning, deep learning, etc., are the branches of artificial intelligence that are showing the phenomenal success that was never seen by humans. In some tasks such as game playing, artificial intelligence has been achieved superhuman performance e[46][33][42].

A couple of the tradeoffs are the interpretability and the availability of computational power. Although there are several works to address this issue, the problems are still continuing and have not been addressed completely [6] [18][44][36][1]. Machine learning, which is a branch of artificial intelligence that learns from data. In broad, there are two types of models in machine learning: generative and discriminative. Generative models [10][47][31][28] [29][22][3][49][32] became very much popular in both research and journalist circles and there is much scope for medical field in this aspect. Although we do not explicitly review the effect of generative models in the health sector,

the review of models done in this paper can be treated as a backbone for such analysis.

Since lot of medical data is available in current days. It is easy to get data from public as well as hospital-collected datasets. Among all other sectors, health sector is a significant and result oriented one at magnificent scale as it has capability or provision of saving many lives as well as allowing a large mass to lead a happy and better life. Among all the diseases and disorders, the autism disorder is also at alarming, especially in India and also in foreign countries. So, we selected the Autism as the baseline disorder and reviewed the potential applications of deep learning models in medical domain.

A. Background

Neural networks are the parametric models in machine learning that are used to solve several types of tasks. Significant tasks that are performed by neural networks are classification, regression, generation, segmentation, etc., Deep neural networks, which are the basis for deep learning are neural networks with more layers. Feature engineering, which is laborious work in traditional machine learning has been replaced by deep neural networks. Each layer contributes its computation to perform the underlying task. Many modern applications such used deep neural networks. There are three popular and important types of neural networks [25], namely, Convolutional neural networks, Recurrent neural networks, Feed-Forward neural networks. Convolutional neural networks have been very successful in tasks that are related to images and videos. Recurrent neural networks in the case of text and time-series data. Feedforward neural networks are used for transforming input into features. Based on the task under consideration, we use the layers of corresponding neural networks that perform well in that particular task.

Since the convolutional neural networks are proved successful in a wide variety of image-related tasks, we discuss in this paper the famous, popular, and useful neural network architectures in literature. Since CNN is a general tool for image-related tasks, it is not restricted to a particular domain. Medicine is one of the domains in which CNN is extensively used. There are many breakthrough performances of CNN in detecting, analyzing, and diagonalizing different types of diseases. CNN is applied in medicine Learning.

Several community-driven or proprietary deep learning packages have been developed by several enthusiastic experts as well as tech giants. TensorFlow, PyTorch, Apache MXNet, Theano, Caffe, fast.ai, The Microsoft Cognitive Toolkit, TFlearn, Lasagne, Elephas, spark-deep-learning, Distributed Keras are the popular packages in contemporary times. In this paper, we discuss the detection of disease using images and CNN.

There are two widely used Convolutional layers in pytorch: Conv2D and Conv3D. Conv2D layer is used in image related tasks and Conv3D is used in video related tasks. Training in deep learning can be simply viewed as an Optimisation problem in mathematics, especially calculus. Along with the model, a deep neural network, there will be a special mathematical function called a loss function which is used to obtain the parameters of the model that fits the underlying data. Each neuron in a neural network consists of a summation unit followed by an activation function. The summation unit applies linear transformation on the input vector using weights and biases.

It can also be viewed as an affine transformation. The purpose of the activation function is to apply a non-linear transformation to the output of the summation unit. If there

are no activation functions in a neural network within the whole output is also a just linear transformation on the input which can be achieved by using a single neuron. Generally used activation functions include Exponential Linear Unit, Hardshrink, Hardsigmoid, HardTanh, hardswish, LeakyReLU, LogSigmoid, MultiheadAttention, PReLU, REctified Linear Unit, ReLU6, Randomized ReLU, SELU, CELU, GELU, Sigmoid, Sigmoid Linear Unit, Mish, Softplus, Softshrink, Softsign, Tanh, Tanhshrink, Threshold, Gated Linear Unit, Softmin, Softmax, Softmax2d, LogSoftmax, AdaptiveLogSoft- maxWithLoss. Among all these ReLU, Sigmoid are popular choices based on whether the task is classification or another. In this paper, we use ReLU activation function in all layers except the last which uses Sigmoid.

Models in machine learning are designs that are useful in obtaining approximate functions. Models are designed by hyperparameters. Models contain parameters. If we consider the Bayesian tree as a model then the number of notes and the number of values taken by a random variable in a node are hyper parameters. Since hyperparameters cannot be learned by a model, they have to have experimented and predicted that designer. Parameters can be learned by the model after enough epochs. Since we are using a deep neural network in our paper, the hyperparameters are are the type of neural network, the number of layers in the neural network, the number of neurons in the each layer, the activation function, batch size, etc., and parameters are weights and biases. After training, we generally get the weights and biases that explain the underlying data. It has been proved that neural networks are capable of approximating any continuous real function.

The objective function is a mathematical function that is highly useful to train the model. Learning parameters can

be viewed as an optimization problem over the objective function. If the objective function is a loss function or cost function then we have to find a gradient Descent algorithm and if the objective function is a performance measure function then we opt for the gradient ascent algorithm. The speed of convergence and several other key factors rely upon the properties of the loss function[43][48][30][8] [39]. For example, if the loss function is convex in nature then the gradient descent algorithm is sure to converge.

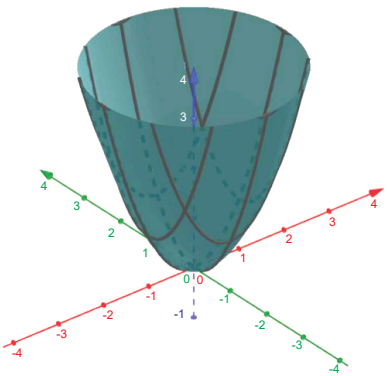


Fig. 1. *Convex structure*

But, unfortunately, the most real-time application needs a loss function that is not convex in nature [9] [37][12]. Differentiability is one of the recommended properties for a function to be a loss function. AS with the case of activation functions, there are several loss functions in literature and many are available in PyTorch. Almost all of them possess useful properties such as differentiability, continuity etc. with some caveats. They include mean absolute error, mean squared error, cross entropy loss, Connectionist Temporal Classification loss, negative log likelihood loss, Negative log likelihood loss with Poisson distribution of target, Gaussian negative log likelihood loss, The Kullback-Leibler divergence loss, Binary Cross Entropy, HuberLoss.

We performed analysis using several loss functions and then presented the one that is giving better results. For simplicity, we xonsiderted the firs three loss functions in the list.

B. Types of Data

The data we are going to consider exists in two forms: labels and images.

In the first type, each example, corresponding to a single autism person contains several fields and each field quantifies a particular aspect of the person such as age, gender, cheek size etc., Feed forward neural networks are apt for this purpose and we have been experimented with the different architectures and the results has been provided.

In the second type, each instance of autism person is an image and image is stored as a tensor in deep learning package PyTorch. This tensor is passed to the CNN and obtain the result that gives the probability of the disorder. Each image is an RGB image and hence each pixel is a triplet. We normalize the image with moving batch normalization.

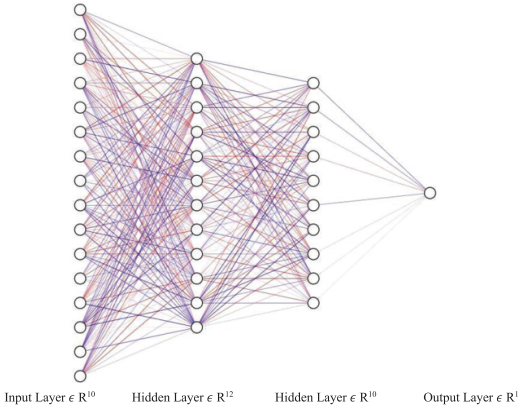


Fig. 2. *Illustration of feedforward neural network*

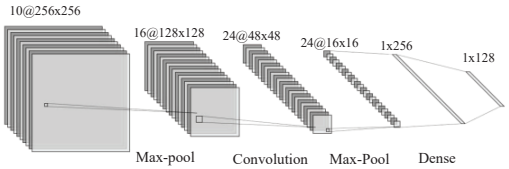


Fig. 3. *Illustration of CNN*

All the popular architectures discussed generally used more layers to minimize the error rare and it does not lead to solution since the vanishing gradient issue is predominant in such cases. Due to the low value of gradients that passes through layers, gradient flow is not as desired. So, we need to address this issue. ResNet came up with the solution. As a result, as the number of layers increases, so does the training and test error rate. ResNet solves the issue by adding the skip connections.

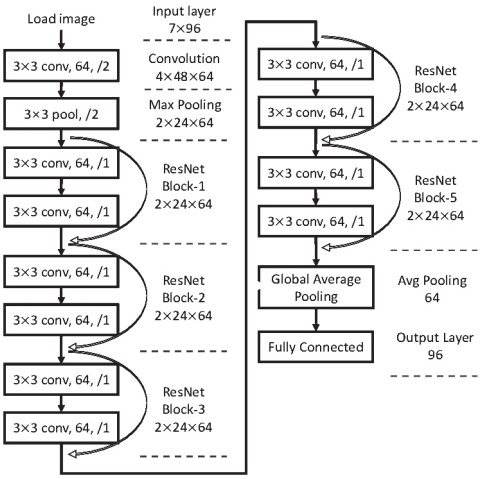


Fig. 4. *Architecture of ResNet*

We use the ResNet architecture for analysis.

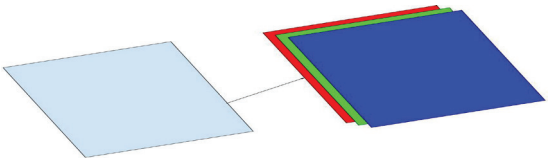


Fig. 5. *Image as RGB tensor in PyTorch*

Figure of CNN shows the intermediate features, which can be called as channels at hidden layers of the CNN.

Architectures Used

We used two architectures at broad level: feed forward neural network, Convolutional neural network.

The progress of the substantial Convolutional architectures roughly starts 1998. LeNet, AlexNet, ZFNet, Inception, VG- GNet, ResNet are the benchmark stages in the architectures of CNN.[26][2][34].

Autism, concerns and data available

Autism[16][14][45] is a major developmental, that cannot be ignored, disability that makes the person to face difficulty in several aspects. It includes difficult to speak and interact with other people in the surroundings such as classmates, teachers. Autism spectrum disorder affects the central neural system, as well as the affected person's general cognitive ability, emotional ability, social ability, and physical health. There are different symptoms for this and they vary based on scope and severity. Difficulty in communication, difficulty in interacting socially, compulsive disorders and repretitive behaviors are the common signs for identifying the victim. Early detection, as well as behavioural, educational, and family therapy, may help to reduce or remove the symptoms and promote proper growth and learning.

Several deep learning approaches have been developed for tasks on autism based on neural networks, ensemble and expert systems in using different sources of data such as in features-based, image, video and timeseries forms [21][50] [38][23][35][27].

In addition to them, there are several parallel or pos- sibly overlapping studies has been conducted on computer vision perspective [19][20][11][5][41][40], which is significant enough to mention as the progress in deep computer vision is phenomenal.

B. Results and Analysis

We performed the initial experiment with a feedforward neural network and then with resnet, the results are as follows.

1) Using feedforward nn:

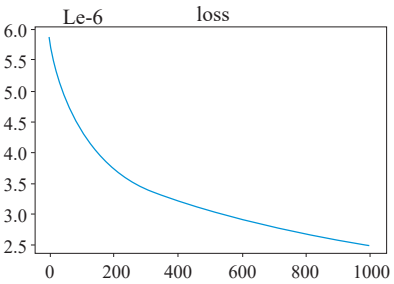


Fig. 6. *feed forward with 15 layers*

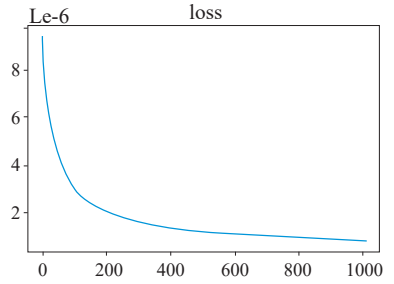


Fig. 7. *feed forward with 20 layers*

2) Using ResNet:

And then we performed the same experiment with resnet and the results are as desired compared to the previous results on naive feed forward neural networks.

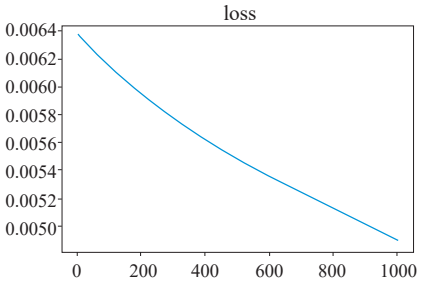


Fig. 8. *ResNet with 10 layers*

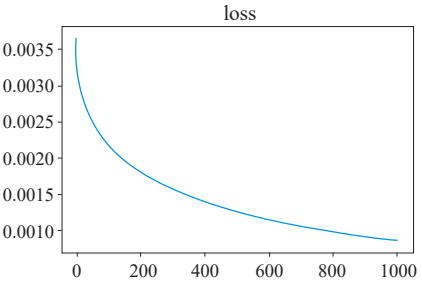


Fig. 9. *resnet with 15 layers*

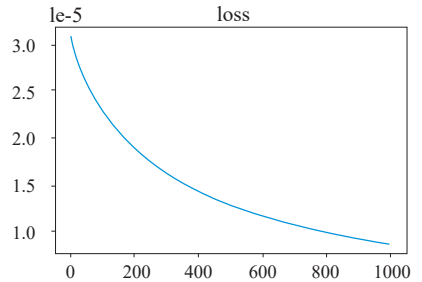


Fig. 10. *ResNet with 50 layers*

As we observe, ResNet overperforms the traditional feedforward neural network. It is due to the resolution of

issue named vanishing gradients in deep learning. In short, the gradients of the loss function approach 0 when more layers with specific activation functions are added to neural networks, making the network difficult to train.

CONCLUSION

We collected the datasets related to autism containing atleast feature-based and image data and then experimented, along with the literature review, with the modern neural networks, including feedforward and Convolutional. The results obtained on autism shows that the deep neural networks will be highly useful for the medical field.

ACKNOWLEDGMENT

We want to thank the University of Hyderabad, for providing the infrastructure to do these experiments, specifically allowing to work in the Artificial Intelligence lab and giving us access to resources. The computational results reported in this work were performed on the resources available in the UOH¿

REFERENCES

[1] Charu C Aggarwal et al. Neural networks and deep learning. *Springer*, 10:978–3, 2018.

[2] Md Zahangir Alom, Tarek M Taha, Christopher Yakopcic, Stefan West- berg, Paheding Sidike, Mst Shamima Nasrin, Brian C Van Esesn, Abdul A S Awwal, and Vijayan K Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*, 2018.

[3] Martin Arjovsky and Le´on Bottou. Towards principled methods for train- ing generative adversarial networks. *arXiv preprint arXiv:1701.04862*, 2017.

[4] Shehan Caldera, Alexander Rassau, and Douglas Chai. Review of deep learning methods in robotic grasp detection. *Multimodal Technologies and Interaction*, 2(3):57, 2018.

[5] Kathleen Campbell, Kimberly LH Carpenter, Jordan Hashemi, Steven Espinosa, Samuel Marsan, Jana Schaich Borg, Zhuoqing Chang, Qiang Qiu, Saritha Vermeer, Elizabeth Adler, et al. Computer vision analysis captures atypical attention in toddlers with autism. *Autism*, 23(3):619– 628, 2019.

[6] Diogo V Carvalho, Eduardo M Pereira, and Jaime S Cardoso. Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8):832, 2019.

[7] Gabriel Chartrand, Phillip M Cheng, Eugene Vorontsov, Michal Drozdzal, Simon Turcotte, Christopher J Pal, Samuel Kadoury, and An Tang. Deep learning: a primer for radiologists. *Radiographics*, 37(7):2113–2131, 2017.

[8] Liangjun Chen, Hua Qu, Jihong Zhao, Badong Chen, and Jose C Principe. Efficient and robust deep learning with correntropy-induced loss function. *Neural Computing and Applications*, 27(4):1019–1031, 2016.

[9] James R Clough, Nicholas Byrne, Ilkay Oksuz, Veronika A Zimmer, Julia A Schnabel, and Andrew P King. A topological loss function for deep-learning based image segmentation using persistent homology. *arXiv preprint arXiv:1910.01877*, 2019.

[10] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1):53–65, 2018.

[11] Ryan Anthony J de Belen, Tomasz Bednarz, Arcot Sowmya, and Dennis Del Favero. Computer vision in autism spectrum disorder research: a systematic review

of published studies from 2009 to 2019. *Translational psychiatry*, 10(1):1–20, 2020.

[12] Jordi de La Torre, Domenec Puig, and Aida Valls. Weighted kappa loss function for multi-class classification of ordinal data in deep learning. *Pattern Recognition Letters*, 105:144–154, 2018.

[13] Li Deng and Dong Yu. Deep learning: methods and applications.

[14] Foundations and trends in signal processing, 7(3–4):197–387, 2014.

[15] Preeti A Devnani and Anaita U Hegde. Autism and sleep disorders.

[16] Journal of pediatric neurosciences, 10(4):304, 2015.

[17] Alessandro Di Nuovo, Daniela Conti, Grazia Trubia, Serafino Buono, and Santo Di Nuovo. Deep learning systems for estimating visual at- tention in robot-assisted therapy of children with autism and intellectual disability. *Robotics*, 7(2):25, 2018.

[18] Hannah Gardener, Donna Spiegelman, and Stephen L Buka. Prenatal risk factors for autism: comprehensive meta-analysis. *The British journal of psychiatry*, 195(1):7–14, 2009.

[19] Fatih Gelgi. Implications of go¨del's incompleteness theorem on ai vs. mind. *NeuroQuantology*, 2(3), 2004.

[20] Leilani H Gilpin, David Bau, Ben Z Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*, pages 80– 89. IEEE, 2018.

[21] Jordan Hashemi, Thiago Vallin Spina, Mariano Tepper, Amy Esler, Vassilios Morellas, Nikolaos Papan-

ikolopoulos, and Guillermo Sapiro. A computer vision approach for the assessment of autism-related behavioral markers. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, pages 1–7. IEEE, 2012.

[22] Jordan Hashemi, Mariano Tepper, Thiago Vallin Spina, Amy Esler, Vassilios Morellas, Nikolaos Papanikolopoulos, Helen Egger, Geraldine Dawson, and Guillermo Sapiro. Computer vision tools for low-cost and noninvasive measurement of autism-related behaviors in infants. *Autism research and treatment*, 2014, 2014.

[23] Anibal So´lon Heinsfeld, Alexandre Rosa Franco, R Cameron Craddock, Augusto Buchweitz, and Felipe Meneguzzi. Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clinical*, 17:16–23, 2018.

[24] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019.

[25] Marjane Khodatars, Afshin Shoeibi, Delaram Sadeghi, Navid Ghaasemi, Mahboobeh Jafari, Parisa Moridian, Ali Khadem, Roohallah Al- izadehsani, Assef Zare, Yinan Kong, et al. Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: a review. *Computers in Biology and Medicine*, 139:104949, 2021.

[26] Joost N Kok, Egbert J Boers, Walter A Kosters, Peter Van der Putten, and Mannes Poel. Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence*, 1:270–299, 2009.

[27] Anders Krogh. What are artificial neural networks? *Nature biotechnol- ogy*, 26(2):195–197, 2008.

[28] Yann LeCun et al. Generalization and network design strategies.

[29] C*onnectionism in perspective*, 19(143-155):18, 1989.

[30] Matthew Leming, Juan Manuel Go´rriz, and John Suckling. Ensemble deep learning on large, mixed-site fmri datasets in autism and other tasks. *International journal of neural systems*, 30(07):2050012, 2020.

[31] Ming-Yu Liu and Oncel Tuzel. Coupled generative adversarial networks.

[32] Advances in neural information processing systems, 29, 2016.

[33] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2794–2802, 2017.

[34] Juan Manuel Martin-Donas, Angel Manuel Gomez, Jose A Gonzalez, and Antonio M Peinado. A deep learning loss function based on the perceptual evaluation of the speech quality. *IEEE Signal processing letters*, 25(11):1680–1684, 2018.

[35] Luke Metz, Ben Poole, David Pfau, and Jascha Sohl-Dickstein. Unrolled generative adversarial networks. *arXiv preprint arXiv:1611.02163*, 2016.

[36] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018.

[37] Thomas M Moerland, Joost Broekens, Aske Plaat, and Catholijn M Jonker. A0c: Alpha zero in continuous action space. *arXiv preprint arXiv:1805.09613*, 2018.

[38] Sanskruti Patel. A comprehensive analysis of convolutional neural network models. *International Journal*

of Advanced Science and Technology, 29(4):771–777, 2020.

[39] Nastaran Mohammadian Rad and Cesare Furlanello. Applying deep learning to stereotypical motor movement detection in autism spectrum disorders. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 1235–1242. IEEE, 2016.

[40] Daniele Rav`ı, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, and Guang-Zhong Yang. Deep learning for health informatics. *IEEE journal of biomedical and health informatics*, 21(1):4–21, 2016.

[41] Divish Rengasamy, Mina Jafari, Benjamin Rothwell, Xin Chen, and Grazziela P Figueredo. Deep learning with dynamically weighted loss function for sensor-based prognostics and health management. *Sensors*, 20(3):723, 2020.

[42] Ognjen Rudovic, Yuria Utsumi, Jaeryoung Lee, Javier Hernandez, Eduardo Castello´ Ferrer, Bjo¨rn Schuller, and Rosalind W Picard. Cul- turenet: A deep learning approach for engagement intensity estimation from face images of children with autism. In *2018 IEEE/RSJ Inter- national Conference on Intelligent Robots and Systems (IROS), pages 339–346. IEEE, 2018.*

[43] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour. Tversky loss function for image segmentation using 3d fully convolu- tional deep networks. In *International workshop on machine learning in medical imaging*, pages 379–387. Springer, 2017.

[44] Manar D Samad, Norou Diawara, Jonna L Bobzien, Cora M Taylor, John W Harrington, and Khan M Iftekharuddin. A pilot study to identify autism related traits in spontaneous facial actions using computer vision. *Research in Autism Spectrum Disorders*, 65:14–24, 2019.

[45] Guillermo Sapiro, Jordan Hashemi, and Geraldine Dawson. Computer vision and behavioral phenotyping: an autism case study. *Current Opinion in Biomedical Engineering*, 9:14–20, 2019.

[46] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.

[47] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 240–248. Springer, 2017.

[48] Neil C Thompson, Kristjan Greenewald, Keeheon Lee, and Gabriel F Manso. The computational limits of deep learning. *arXiv preprint arXiv:2007.05558*, 2020.

[49] Chengzhong Wang, Hua Geng, Weidong Liu, and Guiqin Zhang. Pre- natal, perinatal, and postnatal factors associated with autism: A meta- analysis. *Medicine*, 96(18), 2017.

[50] Fei-Yue Wang, Jun Jason Zhang, Xinhu Zheng, Xiao Wang, Yong Yuan, Xiaoxiao Dai, Jie Zhang, and Liuqing Yang. Where does alphago go: From church-turing thesis to alphago thesis and beyond. *IEEE/CAA Journal of Automatica Sinica*, 3(2):113–120, 2016.

[51] Kunfeng Wang, Chao Gou, Yanjie Duan, Yilun Lin, Xinhu Zheng, and Fei-Yue Wang. Generative adversarial networks: introduction and outlook. *IEEE/CAA Journal of Automatica Sinica*, 4(4):588–598, 2017.

[52] Jingyi Xu, Zilu Zhang, Tal Friedman, Yitao Liang, and Guy Broeck. A semantic loss function for deep learning

with symbolic knowledge. In *International conference on machine learning*, pages 5502–5511. PMLR, 2018.

[53] Xin Yi, Ekta Walia, and Paul Babyn. Generative adversarial network in medical imaging: A review. *Medical image analysis*, 58:101552, 2019.

[54] Jian Zhou, Christopher Y Park, Chandra L Theesfeld, Aaron K Wong, Yuan Yuan, Claudia Scheckel, John J Fak, Julien Funk, Kevin Yao, Yoko Tajima, et al. Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. *Nature genetics*, 51(6):973–980, 2019.

[55] James Zou, Mikael Huss, Abubakar Abid, Pejman Mohammadi, Ali Torkamani, and Amalio Telenti. A primer on deep learning in genomics. *Nature genetics*, 51(1):12–18, 2019.

Modified Convolutional Neural Network Classification of Diabetic Retinopathy

Pankaj Borkar

Department Of Computer Science
Tulsiramji Gaikwad Patil College
Of Engineering And Technology, Nagpur

Anup Gade

Department Of Computer Science
Tulsiramji Gaikwad Patil College
Of Engineering And Technology, Nagpur

Abhimanyu Dutonde

Department Of Computer Science
Tulsiramji Gaikwad Patil College
Of Engineering And Technology, Nagpur

***Abstract*—**Diabetes is a disease that affects the production of blood glucose levels in the body. Uncontrolled blood sugar levels in the body lead to a condition called diabetes. Diabetes can damage several organs in the body, such as blood vessels, the kidney, the heart, and the eyes. Uncontrolled blood sugar levels can lead to a condition

called diabetic retinopathy, or diabetic eye disease. While prior research has looked at the risk factors for diabetic eye disease, it is still unknown which factors are more directly linked to the condition. With the potential ability to distinguish diabetic eye disease-related hazard factors with more exactness, we will be able to develop an early intervention plan of action for diabetic eye disease in those at greater risk. This paper proposes a method for the classification of diabetic eye disease. This investigation is a review of machine learning techniques applied to the detection and classification of images of eyes affected by diabetic eye disease. There are factors that may affect the effectiveness of machine learning algorithms in the classification of eye images affected by diabetes that are also taken into consideration.

Keywords- *Data Mining, Artificial neural fuzzy interference system, K-Nearest-Neighbor (KNN), Machine Learning (ML), Support Vector Machines, Decision Trees*

Introduction

Diabetes is a disease caused by an uncontrolled blood sugar level in the body. Diabetes can lead to several serious health complications, such as damage to the nerves, eyes, kidneys, and heart. This condition has an impact on people of all ages. Non-insulin-dependent diabetes (type-2) and insulin-dependent diabetes (type-1) are two types of diabetes.According to data by the World Diabetes Federation, approximately 423 people are currently suffering from diabetes. In 2018, diabetes caused 1.4 million deaths. And nearly half of all deaths occurred before the age of 65. According to the World Diabetes Federation, diabetic eye disease is one of the major diseases caused by diabetes. Diabetic eye disease occurs as a result of an uncontrolled blood sugar level in the blood, which damages the blood

vessels in the retina. Diabetic eye disease can cause blood vessels to swell. As a result, a person with this disease can experience blurry vision or experience permanent blindness.

A diabetic eye disorder occurs as a result of diabetes. Patients who are diagnosed with diabetes have a tendency to be affected by diabetic eye disease. A person having diabetes for a long period of time is the most likely to get affected by diabetic eye disease. Non-proliferative diabetic eye disease (NPDED) and proliferative diabetic eye disease (PDED) are the two types of diabetic retinopathy that can arise as a result of diabetes. Diabetes damages the retinal capilleries. This damage causes patients to develop neovascularization, microaneurysms, hemorrhages, and hard exudates. The difference between normal and damaged retina is shown in figure 1.

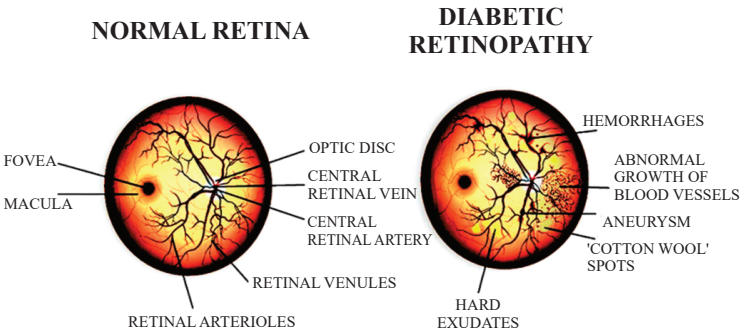


Figure 1: *Difference between normal retina and diabetic retinopathy image*

According to the World Diabetic Federation, early detection of diabetic eye disease is critical in order to take preventive measures. Diabetic eye disease is also difficult to detect early. If the patient with diabetic eye disease is detected early, then serious threats can be avoided. Early

screening can avoid threats such as permanent blindness. The diagnosis of diabetic eye disease using fundus images takes a lot of time. The treatment for diabetic eye disease requires an opthalmologist who is proficient and skilled in the treatment. The self-operating system, which can predict and classify DR images, can be helpful in detecting diabetic eye disease early. It will also save time and money. Moreover, for eyes with clinically significant retina damage, laser photocoagulation needs to be thought about, particularly when the macula's focal point is included or is unavoidably compromised [6]. Treatment for diabetic macular edoema that is antagonistic to vascular endothelial growth factor (against VEGF) has also been established [2].

Color fundus screening is the most effective technique used to detect and classify diabetic eye disease images. But the use of different devices can cause image quality to be destroyed. So, in order to detect and classify diabetic eye disease, the prepossession of images must be done. In order to achieve the correct results, the retinal images must be preprocessed first to enhance the image quality. To get an accurate result, image sharpening algorithms are used to remove noise from the image. Histogram Equalization is the most effective algorithm. This classification model uses a modified convolutional neural network to classify the images as proliferative, normal, severe, mild.

According to a collective investigation of the community-based aggregate research conducted around the sector, diabetes that has been around a long time is identified as a substantial hazard factor. Other risk variables identified in this study included elevated haemoglobin A1C (HbA1C) levels in the fasting state and high blood pressure. According to the study prolonged duration of diabetes increases the risk of diabetic eye disease and controlled blood sugar levels reduce the risk of diabetic eye disease.

In any case, it has been seen in clinical practise that a small number of patients with long-term managed A1c levels are nevertheless at risk of developing diabetic eye disease in non-insulin dependent diabetes.

The findings suggest that the haemoglobin A1C level isn't the most important relevant hazard factor in non-insulin dependent diabetes, and that other factors such as high blood pressure, hyperglycemia, and duration of diabetes may have played only a partial role in the improvement of diabetic retinopathy in T2D. There were discovered to be risk factors for diabetic eye disease, such as a long duration of uncontrolled blood sugar levels, high blood pressure, and high cholesterol, among others.

At the end, the uncontrolled blood sugar level in the body came across as the main contributor to the presence of diabetic eye disease. Poor preventive measures to control sugar in the blood result in an increase in glucose levels in the blood, so the A1c level, pre-meal and post-meal blood sugar increase as well. Every one of the three previously mentioned parameters advises on a different aspect of diabetes and should all be taken into consideration in the treatment of diabetes.

RELATED WORK

A number of different research has been done to find an effective technique to classify diabetic eye disease images. Different methods such as neural networks, SVM, and K-Nearest-Neighbor decision trees are used to differentiate the images from one another. A study of patients was carried out to detect whether patients with diabetic eye disease had non-insulin-dependent diabetes in Iran [11]. The information used in this investigation was collected on the basis of the period of uncontrolled blood sugar level that was taken into consideration. It was discovered that using

computed relapse models with the needed variable (DR) had negative repercussions in this investigation. According to the output, 70% of the patients were affected by the concentration of high blood sugar levels.

In another analysis [12], researchers proposed a model for evaluating a person's health record and the duration of their diabetes. Using patient history can be useful to analyse the severity. The patients' history, type of diabetes, and duration of diabetes are useful to predict the severity of diabetic eye disease. The A1c levels are taken into consideration to predict the risk of diabetic patients developing diabetic eye disease. The AUC was 0.75, suggesting the patients' risk of developing diabetic eye disease is 75 percent higher than previously thought.

The risk of diabetic retinopathy was predicted by Semeraro et al. using the c-measure, the survival beneficiary working trademark, and the Gonen and Heller concordance likelihood gauge (CPE) for the Cox corresponding peril show in a study [13]. The C-file produced an estimation of 0.746 for the inward approval, and the Gonen– Heller CPE for the Cox relative risk approach was 0.683, indicating a reasonable degree of concordance between the observed event of DR and the anticipated event predicted by the model. For the outer approval, the characteristics for C-record and CPE were 0.767 and 0.697, respectively, for the C-record and CPE categories. The AUC for 1-year survival from retinopathy was 0.825, which indicates a high level of accuracy. ($p = 0.137$) There was no statistically significant difference between the C-record of those who were determined in the train informational index and those who were determined in the test informational index. In any case, the classification and relapse tree (CART) examination or the arbitrary woodland investigation are utilised for the train informational collecting in order to confirm that the

outcomes were consistent with the various methods utilised in the investigation.

Information mining approaches were used to create and execute a substructure to calculate the effect of diabetic eye disease, which was developed by Dr. Karim Hashim Al-Saedi and colleagues[14]. In this exploration, the approximation for diabetic eye disease was produced using the information gained from previous studies. The exact and fast methodology of diabetic eye disease prediction gives sufferers enough time to take preventive measures. As a result, the shading fundus image was used to distinguish and comprehend the many sores of diabetic retinopathy and their characteristic characteristics, one by one, in this way. The determinations of the typical shading fundus photographs were split and grouped into two categories: ordinary and weird, using the extraction approach. The abnormal image obtained would then be categorised on the basis of its severity. An association standard and an SVM classifier were used to predict the occult class in order to avoid confusion. The developer predicted positive effects that would benefit both the patients and the speeding up of the procedure.

Abhilash Bhaisare and colleagues [15] have proposed a model to distinguish between the images to identify the DR. The investigation suggested that the picture of the eye's retina can be used to find the various stages of diabetic eye disease. Processing of the image to extract the features of the retina can be done to find out the changes in the retina as compared to the normal retinal image. The obtained result is used to determine whether the image is normal or abnormal to detect the presence of diabetic eye disease.

Naive Bayes and Support Vector Machine computations have been used to predict the early identification of eye disease and DR, according to K. R. Ananthapadmanaban

and G. Parthiban [16]. They discovered that the Naive Bayes algorithm provides 83.37 percent precision while the SVM algorithm provides 64.91 percent exactness when using the Rapid Miner device. The specificity and affectability of these tactics were also evaluated to be 95 percent and 96.65 percent, respectively, for their execution. In the first objective, the flaws created by lightning are removed. In the second objective, the features of the image are preprocessed. For the third objective, the exudation in the image is pointed out. Finally, to summarise, the neural system produces a better output by taking out features in the image. The most recent outcomes were compared numerically, and a manual exudates division was made by an eye professional to help with the comparison.

Three-arrange framework for early detection of Microaneurysm utilising channel banks has been proposed by M. Usman Akram, Shehzad Khalid, and Shoab A. Khan [17] in an article published in the journal NeuroImage. The framework extricates all possible candidate districts for Microaneurysm that may be present in the retinal image using the information provided. In light of several qualities, for example, shape and shading; power; and measurements, an element vector for each area is constructed to classify a hopeful district as either a Microaneurysm or a non-Microaneurysm (see figure). It is demonstrated how a crossover classifier, which combines the Gaussian blend show (GMM), the support vector machine (SVM), and the extension of multi-display medoid-based showing approach, may improve the precision of classification in a gathering.

The real Microaneurysm districts are selected and described using a hybrid classifier, which is a weighted combination of multivariate m-Mediods, GMM, and SVM, among other techniques. Researchers Vimala Balakrishnan

and colleagues [18] have dealt with the topic of integrating association recommendations and case-based reasoning to predict the development of retinopathy. Using information mining, namely association rules using Apriori calculation, and case-based thinking to predict retinopathy, they have presented a framework for predicting retinopathy expectations. Case-based thinking is used to recover comparable cases, while association rules are used to break down examples in the informational index and to determine the likelihood of retinopathy developing. With the utilisation of this framework, the inventor believed that remarkable enhancements may be provided to medical professionals and diabetics alike.

An overview of automatic recognition of non-proliferative DR has been provided by M.Tamilarasi and Dr.K.Duraiswamy [19], respectively. Because there are numerous information mining strategies and calculations that can be used to analyse DR in retinal fundus images, the creators of this paper have checked on, characterised, and examined the calculations and methods that have recently been proposed in order to develop better and more powerful calculations going forward.

Ramon Casanova and colleagues [20] focused on the use of Random Forest (RF) algorithms to DR classification investigations that were based on fundus photography information. An approach to dealing with DR hazard evaluation that relies on metrics obtained from reviewed fundus photography and fundamental information was proposed by the authors. They claimed that radio frequency (RF) could be a valuable instrument for analysing DR findings and assessing their movement.

In order to predict the hazard factor of DR, S.Sagar Imambi and T.Sudha [21] have coupled content mining classification systems on the available clinical information.

In their study, they demonstrated that a reasonably brief period of time in which the executives commenced the case before the commencement of therapeutically observable diabetic eye disease reduced the risk of developing diabetic eye disease in sufferers with non-insulin-dependent diabetes by nearly twofold. In their study, they demonstrated that a reasonably brief period of time in which the executives commenced the case before the commencement of therapeutically observable diabetic eye disease reduced the risk of developing diabetic eye disease in sufferers with non-insulin-dependent diabetes by nearly twofold. The target was to design a grading method that could be available to a large number of ophthalmologists to encourage them to treat patients successfully in order to reduce the hazards of diabetic eye disease related disorders.

To construct a prophetic framework for DR, Hayrettin Evirgen and Menduh erkezi [22] used Naive Bayes classification calculation to investigate the real-world information and come up with a prescient framework. When creating the expectation framework, a total of 385 diabetic patients' information was used. With cross-approval, the researchers revealed that gullible Bayes calculation can be used to forecast diabetic retinopathy with an improved precision of 89 percent.

prediction of diabetic retinopathy using different algorithms

The anticipation of DR can be thought of as a multi-class classification problem. In order to anticipate DR, we consolidated decision trees, calculated relapse, created fake neural systems, and bolstered vector machines. To create expectation models, the SAS Enterprise Miner 12.1 programming was used in conjunction with SAS Enterprise Miner 12.1 programming.

We used decision trees to construct interpretable recommendations for clinical practise to assist with the development of basic leadership forms. We create choice trees and rules for clinical basic leadership, arranging based on information gathering and unbiased assessment, and creating choice trees as foreseeing models to aid in clinical basic leadership. We also conduct clinical basic leadership training.

Using probability distributions, calculated relapse estimates the relationship between the absolute ward variable and at least one autonomous component in a given situation. The first anticipates a calculated capacity, whereas the second anticipates a standard, typical dissemination task to be completed. If the dependent variable meets the criteria of a case, the likelihood of the dependent variable meeting those criteria is proportional to the exponential capacity of the direct relapse mechanism. This demonstrates how the logic fills in as a connecting work between the likelihood of relapse and the direct relapse articulation of the problem. In addition, we included stepwise selection to select discriminative highlights in the context of strategic relapse.

Insights learning hypothesis is a machine learning calculation proposed by Vapnik that is based on the basic hazard minimization rule of the insights learning hypothesis. It is possible that it will be used to deal with concerns such as classification and relapse. Because diabetic retinopathy forecasting is a paired classification problem, support vector machines (SVM) would be beneficial for our motivation. Model improvement is accomplished through the use of spread premise work (RBF), which is the bit work in SVM throughout this time period.

In computing, an artificial neural network (ANN) is a collection of quantifiable learning models pushed by organic neural systems that are used to evaluate or inexact

capacities that can be based on a large number of variables and are frequently obscure. ANNs are typically depicted as frameworks of linked neurons that communicate with one another through the use of signals. Positioned forecast execution of choice trees based on a 60-month, 120-month, or 180-month diabetes term can be changed based on experience, making ANN adaptable to different sources of information and well-suited for machine learning applications. Figure 2 shows the classification using a convolutional neural network. Figure 3 shows the images of different stages of diabetic eye disease.

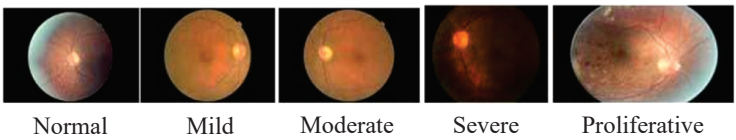


Figure 2*: Convolutional neural network classification process*

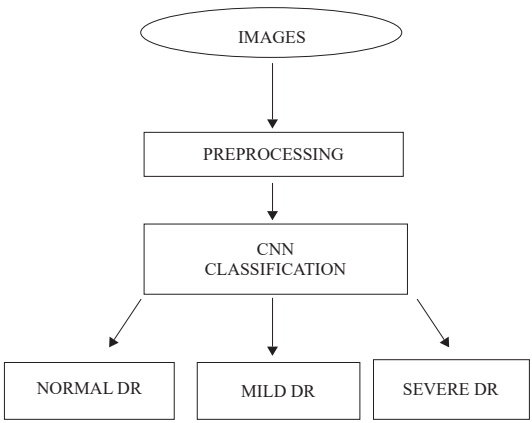


Figure 3*: Different stages of diabetic eye disease*

RESULTS AND ANALYSIS

This research looked at machine learning techniques and deep training procedures for categorising lesion pictures,

combining CNN with other pre-trained models. The advantages and disadvantages of each algorithm Choosing the correct classification approach is crucial to getting the best results. However, CNN outperforms ML approaches that use pre-trained models such as AlexNet, VGG16, and ResNet. The following table gives the accuracy chart for different algorithms.

TABLE 1:Different algorithms and their accuracy

ALGORITHM	ACCURACY (%)
CNN	92
SVM	90
Naïve Bayes	87
kNN	81
Logistic Regression	79
Decision Tree	66

Conclusion

The proximity of haemorrhages and microaneurysms in endoscopic photographs is important in predicting and determining the presence of Diabetic Retinopathy. In order to programmatically recognise diabetic retinopathy via highlight extraction, a plethora of calculations have been developed and implemented. An overview of numerous information mining methods is presented in this research, which reveals that KNN and SVM have the highest correctness when compared to the other approaches. This survey paper can serve as a valuable resource for future scientists who are interested in forecasting diabetic retinopathy through the use of information mining procedures. This will also be beneficial for scientists in order to gain a better understanding of this area in order to make more accurate calculations in the future.

References

[1] Chew EY, Klein ML, Ferris FL 3rd, Remaley NA, Murphy RP, Chantry K, Hoogwerf BJ, Miller D., "Association of elevated serum lipid levels with retinal hard exudate in diabetic retinopathy. Early treatment diabetic retinopathy study (ETDRS) report", 22. Arch Ophthalmol. 2016; 114(9):1079–84.

[2] American Diabetes Association. Standards of medical care in diabetes–2014. Diabetes Care. 2014;37(Suppl 1):S14–80.

[3] Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, Ferris FL 3rd, Klein R, "American Diabetes A. Retinopathy in diabetes. Diabetes Care", 2004;27(Suppl 1):S84–7.

[4] Kempen JH, O'Colmain BJ, Leske MC, Haffner SM, Klein R, Moss SE, Taylor HR, Hamman RF. "The prevalence of diabetic retinopathy among adults in the United States", Arch Ophthalmol. 2004; 122(4):552–63.

[5] Huang YY, Lin KD, Jiang YD, Chang CH, Chung CH, Chuang LM, Tai TY, Ho LT, Shin SJ. "Diabetes-related kidney, eye, and foot disease in Taiwan: an analysis of the nation wide data for 2000-2009", J Formos Med Assoc. 2012; 111(11):637–44.

[6] Early Treatment Diabetic Retinopathy Study Research Group. "Focal Photocoagulation treatment of diabetic macular edema. Relationship of treatment effect to fluorescein angiographic and other retinal characteristics at baseline: ETDRS report no. 19", Early treatment diabetic retinopathy study research group. Arch Ophthalmol. 1995; 113(9):1144–55.

[7] Chang TJ, Jiang YD, Chang CH, Chung CH, Yu NC, Chuang LM. Accountability, utilization and providers for diabetes management in Taiwan, 2000–2009: an

analysis of the National Health Insurance database. J Formos Med Assoc. 2012;111(11):605–16. Tsao et al. BMC Bioinformatics 2018, 19(Suppl 9):283 Page 120 of 121

[8] Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, Chen SJ, Dekker JM, Fletcher A, Grauslund J, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012; 35(3):556–64.

[9] UK Pospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. UK Prospective Diabetes Study Group. BMJ. 1998; 317(7160):703–13.

[10] Kowall B, Rathmann W. HbA for diagnosis of type 2 diabetes. Is there an optimal cut point to assess high risk of diabetes complications, and how well does the 6.5% cut-off perform? Diabetes Metab Syndr Obes. 2013; 6:477–91.

[11] Hosseini SM, Maracy MR, Amini M, Baradaran HR. A risk score development for diabetic retinopathy screening in Isfahan-Iran. J Res Med Sci. 2009;14(2):105–10.

[12] Aspelund T, Thornorisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, Einarsson S, Palsson O, Einarsson G, Bek T, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525–32.

[13] Semeraro F, Parrinello G, Cancarini A, Pasquini L, Zarra E, Cimino A, Cancarini G, Valentini U, Costagliola C. Predicting the risk of diabetic retinopathy in type 2 diabetic patients. J Diabetes Complicat. 2011;25(5):292–7.

[14] Dr. Karim Hashim Al-Saedi, Dr. Razi Jabur Al-Azawi, Rasha Asaad Kamil, - Design and Implementation Sys-

tem to Measure the Impact of Diabetic Retinopathy Using Data Mining Techniques, International Journal of Innovative Research in Electronics and Communications (IJIREC) Volume 4, Issue 1, 2017, PP 1-6

[15] Abhilash Bhaisare, Sagar Lachure, Amol Bhagat, Jaykumar Lachure - Diabetic Retinopathy Diagnosis Using Image Mining, International Research Journal of Engineering and Technology (IRJET), Volume: 03, Issue: 10, Oct -2016

[16] K. R. Ananthapadmanaban and G. Parthiban. -Prediction of Chances - Diabetic Retinopathy using Data Mining Classification Techniques. Indian Journal of Science and Technology, Vol 7(10), 1498–1503, October 2014

[17] M. Usman Akram, Shehzad Khalid, Shoab A. Khan - Identification and classification of microaneurysms for early detection of diabetic retinopathy, Pattern Recognition, Vol. 46, No. 1, 2013, pp. 107–116.

[18] Vimala Balakrishnan , Mohammad R. Shakouri and Hooman Hoodeh - Integrating association rules and case-based reasoning to predict retinopathy, Maejo Int. J. Sci. Technol. 2012, 6(03), 334-343, ISSN 1905-7873

[19] M. Tamilarasi and Dr. K. Duraiswamy -A Survey for Automatic Detection of Non- Proliferative Diabetic Retinopathy, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2014

[20] Ramon Casanova , Santiago Saldana, Emily Y. Chew, Ronald P. Danis, Craig M. Greven, Walter T. Ambrosius - Application of Random Forests Methods to Diabetic Retinopathy Classification Analyses, Published: June 18, 2014https://doi.org/10.1371/journal.pone.0098587

[21] S. Sagar Imambi and T. Sudha - Building Classification System to Predict Risk factors of Diabetic Retinopathy Using Text mining, (IJCSE) International Journal

on Computer Science and Engineering Vol. 02, No. 07, 2010, 2309-2312

[22] Hayrettin Evirgen, Menduh Çerkezi- Prediction and Diagnosis of Diabetic Retinopathy using Data Mining Technique,ijaedu.ocerintjournals.org/tojsat/issue/22642/241874, Sep, 2014

Learning Education Using Artificial Intelligence

Mrs.R.S.Archana Vishveswari

Assistant Professor,
Department of Computer Science and Engineering,
Varuvanvadivelan Institute of Technology,
Dharmapuri, TamilNadu

Dr.Senbagavalli M

Associate Professor,
Department of Information Technology,
Alliance College of Engineering & Design,
Alliance University, Bangalore

Abstract- Artificial Intelligence on education gives a detailed architecture and framework for assess Artificial Intelligence from a preliminary analysis. In this field of study about the field of new innovations and the developments that have included in machine Learning and other artefacts having Artificial Intelligence characterized by the following capabilities like abilities, learning, adaptability, and decision-making. This paper was to know about the impact of Artificial Intelligence in education

sector through learning activities. In this research approach, Artificial Intelligence shows how to mind a super computer, including adaptive behaviour, how the sensors, and other capabilities, that enable to behave like have human like and functional abilities, and indeed, It improves the super computers interaction between the human begins in the education sector, it has been increased application of Artificial Intelligence, going over the understanding of Artificial Intelligence as a super computer to include computer systems. An Artificial Intelligence opens a new capabilities and possibilities in education sector, where districts leaders can get more and more insight into how the Artificial Intelligence will enhance the teaching sector in future**.**

Keywords: *Artificial Intelligence. Decision-making, cognitive abilities, machine learning, enhance teaching, cognitive learning.*

I. INTRODUCTION

Artificial Intelligence will injects more complex and collaborative learning activities in education sector along with the students knowledge and their assessments, In Artificial Intelligence and the future of Artificial Intelligence, a detail analysis from digital platform and the centre for research in computing and learning sciences. Sometimes, it is necessary to research beyond the limits and the norms, to develop and introduce a different ways of teaching abilities using Artificial Intelligence. In the introduction of Artificial Intelligence, in advancements, and proliferation, AI, has made more simple to learn and move forward more effectively and efficiently. Before entering into the paper we should come to know how the computers and other technologies, instructs our educators and students, to learn

through learning technology by the complete involvement of human developers.

Use of Artificial Intelligence in education sector plays a major role and it creates a great impact, and it includes improvement in learning, smarter thinking like a human and improved learning capability, effectively and efficiently in education administration along with other sectors. Artificial Intelligence continuously developed, and it also produces a new ways of application in education emergency day by day.

A. Artificial Intelligence in current Education

Now a days there is a rapid development in education field, Artificial Intelligence (AI) and machine learning (ML), robotics in all organisation including education field. If the education field starts to utilize the Artificial Intelligence in full potential in everywhere, the upcoming next generation students will use Artificial Intelligence through media and there learning level increases the learning capabilities in educational application.

In this research approach, Artificial Intelligence shows how to mind a super computer, including adaptive behaviour, how the sensors, behave like human, and indeed, It improves the super computers interaction between the human begins in the education field, it has been increased the scope of Artificial Intelligence, by understanding the Artificial Intelligence as a super computer to include computer systems., a computer has a massive processing capabilities, which includes an adaptive behaviour, such as encompassing of sensors, and other capabilities in the education field, also there has been a higher level practice of Artificial Intelligence, grows higher and higher, the understanding of Artificial Intelligence as a supercomputer and it includes embedded computer systems.

The application of Artificial Intelligence and machine learning techniques in education field are creating interest year by year. It is a branch of science producing and studying the Machines Learning at the incentive of human intelligence processes.

Artificial Intelligence mostly comes across these three basic principles:

1. Learning
2. Self_Correction
3. Reasoning

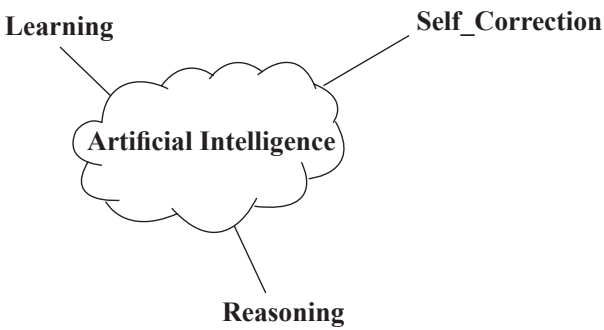


Figure1: *Basic Principles of AI*

1. Learning:

Learning is a process of asking questions like where, when, what, why, etc., When we start asking why there our knowledge improves. Likewise in Artificial Intelligence to asking question makes the students to improves there learning capability.

Learning is an Art which uses students to ask more repeated questions to the instructors, about what they learned, and it makes the students learning capability effectively. Also the human brain always ask questions when

they are learning a things, so this capability of questioning makes the student more powerful in their application.

2. Self_Correction:

By learning the concepts they can easily corrected themselves by their own. when the child strarts to read their own, we should make sure that they completely understand the application or the text or we should know that they can able to clarify their misunderstandings or words they read. Normally in this generation students they are in very emerg situation to enhance their reading capabilities. Reading without any misunderstanding is a very big task which we have in front of our students.

Now a days in Internet we have so many correcting tools to develop our student reading capabilities, like looking or watching whether their sounds for the words are correct or not. So by using such AI tools we can assure that they read correctly. Also they can enhance the reading capabilities.

3.. Reasoning:

Artificial Intelligence helps the children's to find the reasons for all like why and where. Its keep their knowledge to upgrade day by day. So they can find the answers for all the situations without the need of the third persons.

B. Artificial Intelligence and the future education field

In current education scenarios most of the schools in foreign countries took a bold and daring decision to implement Artificial Intelligence, they implemented and succeeded too. These Machine learning and Artificial Intelligence tools improves the student learning capabilities and makes more involvement on their studies through

virtually. Virtual Reality which makes the every student to concentrate.

These techniques can change the learning experience, and changes the experience of virtual learning capability higher than the existed system. Here there is no need of instructor to teach them line by line only they observes whether there are learning their subjects without any doubts. This Artificial Intelligence and Machine Learning technology makes huge change in the entire educational sectors. In addition, these learning tools will allow the students to make their own educational projects even they are not in the place, these systems makes the real-time learning efficiently and makes learning possible.

There is more benefits while we makes the education sector under the Artificial Intelligence, students like physically challenged can also learn their subjects from their own place without the tutors. This makes education sector more comfortable to average students too. The benefits of implementing an artificial intelligence in the classroom, the developments of the students increased day by day. Here it comes some of the possible capabilities.

- Personalizing learning techniques
- Automation of learning techniques
- Adapting a new learning techniques

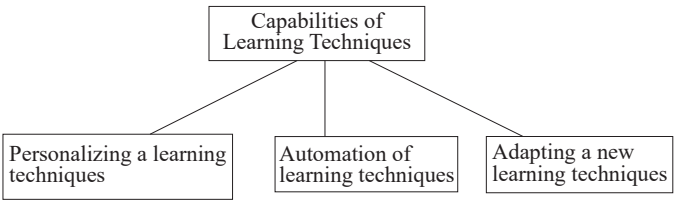


Figure 2: *Capabilities of AI*

1. Personalizing learning techniques:

In this Personalizing learning techniques we can develop our own application as their need of the educational institution.

- So we can develop the learning capabilities as per the knowledge of the children's.
- It helps the students to clarify the doubts in their language.

2. Automation of Learning Techniques:

Here in these techniques students get the clarification about the information what they want.

- Just by typing the word what they want to know, they can come to know, through the speech recognition software they can clearly come to know the concept and also by the virtual reality.

3. Adapting a new Techniques:

By adapting a new technique they can improve their knowledge and the level of confidence. So here by learning a new technique they can compete with the growing world boldly.

- In a growing world updating a new technology makes the students more confident and they can adapt anywhere without any hesitation.

C. LONG-TERM GOAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION FEILD

In future Artificial Intelligence can replace the role of a teacher? This is the question among most of the educators. But seriously I won't replace the role of a teacher but it can reduce the load of the teacher. Particularly while in

lockdown educators teachers are fed up with the online class. Here in this situation Artificial Intelligence helped educators to teach students.

"Edtech company Promethean surveyed teachers and learned that 86 percent thought AI should be an important part of education". AI thinks like a human, Acted like a human, so in future AI plays a main vital role in education sector. By using AI in education field in future students can learn through machine learning and they can learn the things without help of the tutors may or may not be in online.

Advantage of learning environment by using AI can adapt to change the entire environment of the education field. They can able to adapt the new environment quickly.

II. current Education Sector and the Role of Artificial Intelligence

In current education sector, role of AI plays a key role. It's a part of science, developed by the human beings for their self improvement. The important summary of AI is to make the human burden little smoother than they do harder.

Machines make the work smarter but not more than the human being. Machines also need a man power to maintain them. In older days we need a large ledger to maintain all records. But now the computer maintains the ledgers but in there also man power makes the machine to maintain the data.

The role of Ai varies in different platforms like digital marketing, real estate, education sector, private organization etc.,. Now a days schools and colleges are using AI to increase their level of teaching ability. AI helps instructors to develop the ability much more than the before so the students performance level also increases and the result will be seems too good, when compare to the before progress.

Here the teachers will have more time to concentrate on developing their knowledge as well as students to.

III PLANNING AND ORGANIZING AN ARTIFICIAL INTELLIGENCE

Planning and organizing can occur day by day in this computer generation only the way of planning will be differ. Organizing is a process that maintains or make possibility to lineup the AI conflicts. Each and every organization has a different types of planning strategies. Each strategies plays a different roles. Analyzing process alone changes but the solution for the conflicts will not be change.

IV. IMPLEMENTING STATEGY MODEL

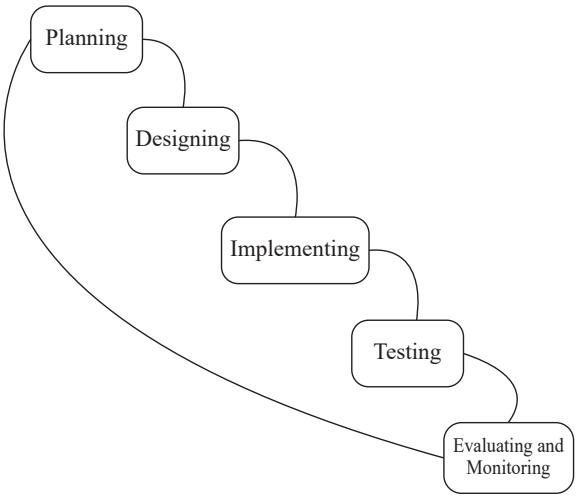


Figure 3: *Implementing Strategy Model*

1. Planning:

Planning is a process which gives a detailed structure of what we need. It makes the entire design to understand neatly. Through Planning the requirements of the process

will come to know clearly. So anyone can easily understand the process.

2. Designing:

Designing gives the overall picture of the about the planning. So we can see the entire process by a detailed view. Designing makes the entire system easily to understand.

3. Implementing:

Through implementing the designing process the overall structure of the planning will be executed as per the planning and designing module. Implementing a process gives the overall body structure for the processed planning architecture

4. Testing:

It aims to ensure that the actual software is free of defects, and to check and determine whether it matches the expected requirements, as designed and planned in the first two process.

5. Evaluating:

Evaluation is the last process which gives the overall result of the process and made the process possible . if test result cause negative then the process should be started again from the first phase, but here we don't need to analyze from the first , only in which module we have an issue we analyze from the particular phase. Again the same process will be executed until the result will be positive.

VI. CONCLUSION:

Here we come with the conclusion in future AI plays a vital role in education sector by implementing a Machine Learning Techniques. Here we should clearly come to know that the machine is not an alternate for tutors but by implanting a machine with the tutor's education sector and

the students they can easily adapt the knowledge effectively and efficiently. But implementing a Machine Learning means doesn't mean to be alternate for tutors its only can reduce a load of the tutors. AI won't change the entire education sector but it changes the way of instructing.

VII. REFERENCES

[1] https://medium.com/towards-artificial-intelligence/artificial-intelligence-in-education-benefits-challenges-and-use-cases-b52d8921f7a

[2] https://link.springer.com/content/pdf/10.1186%2Fs41239-19-0171-0.pdf

[3] https://ieeexplore.ieee.org/document/90698 75

[4] Bullock, L. The Top 6 Ways That Artificial Intelligence Will Affect Your Business In The Near Future. Forbes

[5] Maga- zine. Kim, M. J., Kim, K. J., Kim, S., & Dey, A. K. (2018). Performance evaluation gaps in a real-time strategy game between human and artificial intelligence players. IEEE Access, 6, 13575-13586.

[6] Hernes, M. (2014). A cognitive integrated management support system for enterprises. In International Conference on Computational Collective Intelligence (pp. 252-261). Springer, Cham.

[7] Hernes, M., & Sobieska-Karpińska, J. (2016). Application of the consensus method in a multiagent financial decision support system. Information Systems and e-Business Management, 14(1), 167-185.

[8] Nieto, Y., Gacía-Díaz, V., Montenegro, C., González, C. C., & Crespo, R. G. (2019). Usage of machine learning for strategic decision making at higher educational institutions. IEEE Access, 7, 75007-75017.

[9] https://www.ieee.org

[10] https://iaied.org/about

Recurrent Neural Network – An overview and applications in Medical Science

Kamran

Alliance College of Engineering and Design,
Alliance University, Bangalore, India

Shobana Padmanabhan

Associate Professor
Alliance College of Engineering and Design
Bangalore, India

***Abstract*—**Recurrent neural network or RNN is among the most sought-after neural network architectures in the Artificial neural network sphere and is used across numerous problem statements involving deep learning. This study is aimed at describing the architecture and working of RNNs and its application in the field of medicine and medical practice and how deep learning is enabling the medical science to be tailored to an individual person of interest.

Keywords*—RNN, Deep Learning, Application of RNN in medical science, LSTM, GRU, Bidirectional RNN.*

INTRODUCTION

Recurrent neural networks (RNNs) are a subdivision of artificial neural networks which works on sequential or time series data. Its looped architecture enables it be able to record past data points and outputs to form a following point or new output based on it. [1]

RNN enables other prominent architectures such as LSTM (Long short-term memory) and GRU (Gated recurrent units). The general working flow of the architecture being common in all the variant, where the output of the previous layer feeds into the next input layer. [2]

- Recurrent neural networks can be used for various purpose but the most sought-after application is sequence generation. Along with it, the architecture is capable of performing the below mentioned and explained actions.
- Sentence or sequence generation based on the previous output.
- Sequence classification
- Labelling or identifying and masking of sensitive information (NER). [3]
- The RNN are further classified into different kinds, or to put it simply, the architecture enabled the creation of newer architectures which are namely;

  Bidirectional recurrent neural network (BRNN): As we now know that the output of the previous layer feeds as an input to the next. In BRNN, the layer after the current can also be used to better the accuracy of the output. Hence, bidirectional.

  Long short-term memory (LSTM): One of the most sought-after variation of RNN, LSTM was presented

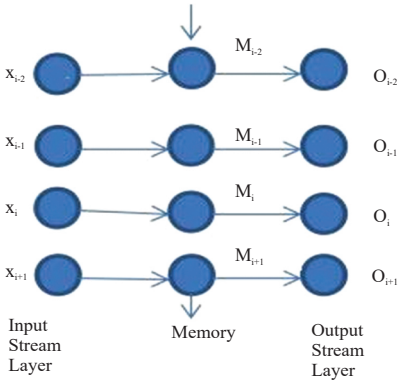
by "Sepp Hochreiter and Jürgen Schmidhuber" addressing the vanishing and exploding gradient problem. LSTM was designed to overcome the issues being faced by the vanilla model by creating a bridge to overcome the gradient problem.

Gated recurrent units (GRU): Similar to LSTM, GRU has two gates, Reset and Update which helps in deciding and retaining information(output) for future outputs or predictions. [4]

Architecture

Recently, the advent of deep learning and recurrent neural networks (RNNs) has led to numerous successes in an array of natural language processing applications like voice, sentiment recognition. Deep learning methods have proven to be able in automating feature-representation and are gradually attempting to eliminate the tedious task of manual intervention.

Recurrent neural networks are used in sequence analysis to figure out the intent by defining dependencies among time stamps. RNN consists of recurrent layers, which are sequentially designed to effectively map the pattern with other sequences and has a strong capability in capturing the contextual data from the sequence. RNN can operate the sequences up to a small length. [5]



*(***Fig. 1** - *Generalised recurrent neural network)*

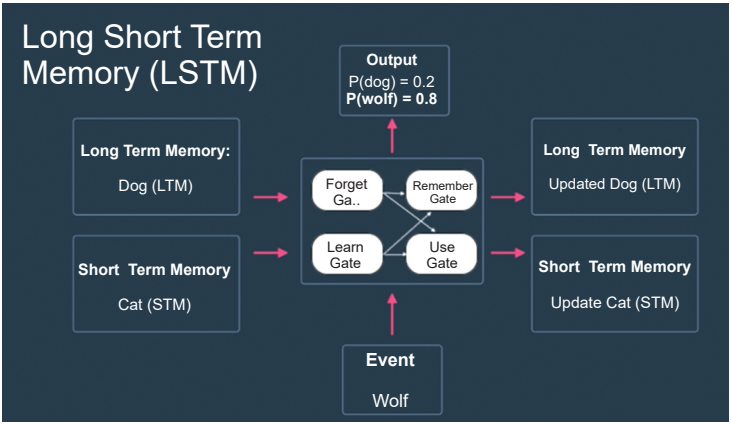
RNN is an addon of feedforward neural network, having loops in the hidden layers. It accepts an input with an output sequence and establishes and identifies the time relationship in-between, helping LSTM in classifying issues. This is done by addition of the network parameters with the hidden node and releasing the state, based on the input values. The vanilla RNN node has single bias and weight and is evaluated by the gated recurrent unit and long short-term memory. On the contrary, LSTM consists of four bias or weights as specified below:

- Forget gate
- Input gate
- Output gate
- State gate

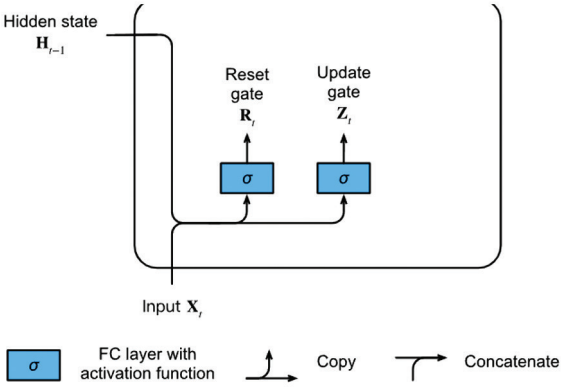
In addition to it GRU consists of only two gates, namely, Reset and Update layer.

"In LSTM, the input gate, output gate, and the forget gate activation is scaled using sigmoid function, and the hidden layer output is filtered using a hyperbolic function. The optimization of network parameters using the stochastic

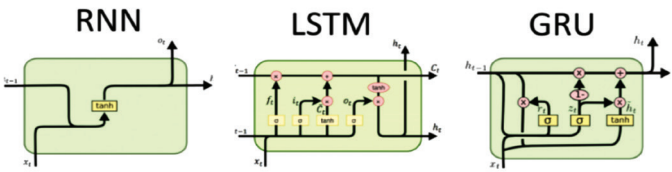
gradient is achieved based on the sequence of input data." [6][7][8]



*(**Fig. 2** - LSTM Architecture)*



*(**Fig. 3** - GRU Architecture)*



(Fig. 4 - *RNN, LSTM and GRU Comparison)*

Applications in medical science

The ability of RNN and its variants as LSTM, GRU enables the architecture to excel in cases of identifying similar conditions as in the case of classification and also in terms of finding the next possible output by prediction. Such ability is widely bartered and used in the medical science sphere, some of which are discussed below to show more insight on their working and application in the said studies.

1. "Survey on RNN and CRF models for de-identification of medical free text

Survey on RNN and CRF models for de-identification of medical free text.

By: Joffrey L. Leevy, Taghi M. Khoshgoftaar & Flavio Villanustre

Journal of Big Data volume 7, Article number: 73 (2020)"

The study points out to the privacy issue of health care data as most of the data is electrically saved without any proper security.

The authors have talked about the feature of the RNN architecture which lets it to recognize sensitive or relevant data by sequence identification or labelling, in layman, to identify and mask the sensitive information about every patient to prevent any abuse of information.

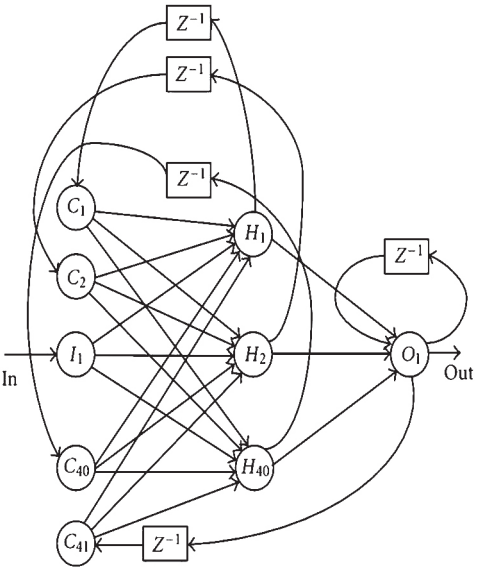
The go-to name for this particular problem and solution is NER or named entity recognition, where the system learns and identifies the respective data which is further masked from general access to prevent any leakage. [9]

1. "Recurrent Neural Networks in Medical Data Analysis and Classifications

Chapter 7 – Applied Computing in Medicine and Health by Al-Jumeily, Dhiya Hussain, Abir Mallucci, Conor Oliver, Carol"

The chapter talks in depth about RNN and how its application in medical science is helpful. One such paper discussed here is **titled "Recurrent neural networks in computer-based clinical decision support for laryngopathies: an experimental study By: Jarosław Szkoła, Krzysztof Pancerz, Jan Warchoł"** stating that RNN architecture has been exceptional in identifying patterns in speech to identify cases of laryngopathy.

The algorithm used here for the purpose is another variation of RNN, namely, the Elman-Jordan neural network. [10]



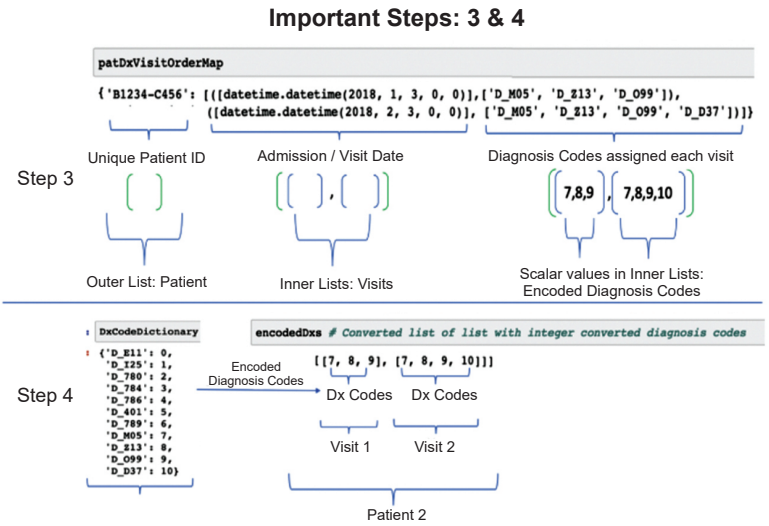
*(***Fig. 5** - *Elman-Jordan Algorithm)*

Input: S—a speech signal of a given patient (a vector of samples), N—a neural network.
Output: $\overline{E}_N$—an average mean squared error corresponding to deformations in S.
$W_{all} \leftarrow Div2Win(S)$;
$W_{sel} \leftarrow SelWin(W_{all})$;
for *each window* $w \in W_{sel}$ **do**
  $Train(N,w)$;
  for *each window* $w^* \in W_{sel}$ **do**
    if $w^* \neq w$ **then**
      $E[w^*] \leftarrow MSE(Test(N,w^*))$;
    end
  end
  $\overline{E}[w] \leftarrow Avg(E)$;
end
$\overline{E}_N \leftarrow Avg(\overline{E})$;
Return $\overline{E}_N$;

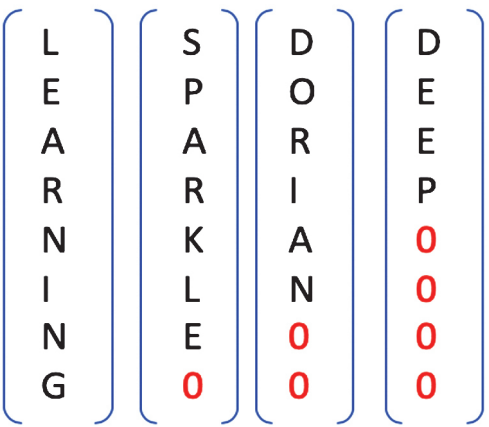
*(***Fig. 6 -** *"Algorithm for calculating an average mean squared error corresponding to deformations in a speech signal.")*

2. ***"Predicting future medical diagnoses with RNNs using Fast AI API:"***

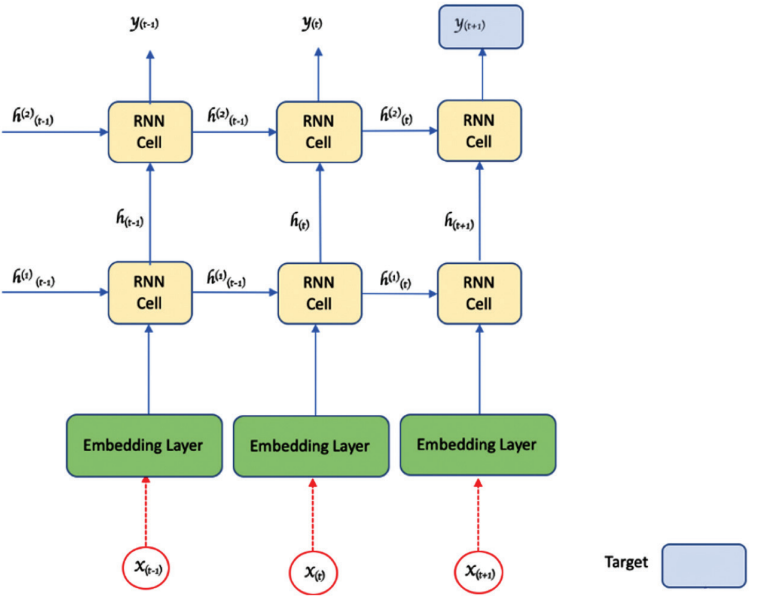
A recreation of the paper "Doctor AI" by Edward Choi. The study uses MIMIC III health dataset, which is an intensive care unit dataset and using the GRU architecture which performs better than LSTM here as done in the original paper. The aim here of the architecture is to predict (all) the diagnosis and medication categories for a subsequent visit. [11]



*(**Fig. 7** - Patient Sequence Encoding)*



*(**Fig. 8** - Variable Length Sequence Padding)*



*(***Fig. 9** - *The model architecture)*

3. "Improved Classification of Brian Tumor in MR images using RNN Classification Framework:"

The paper discusses a straight forward classification architecture divided in three phases with the core being an RNN architecture detecting the regions of possible brain tumor after getting a segmented region by using K means algorithm. [12]

The algorithm as stated:

Step 1: Reading the MRI image

Step 2: Improving the contrast of the MRI

Step 3: Converting the then MRI binary image into grayscale

Step 4: Segmenting the grayscale input using K-Means clustering

Step 5: Extracting features from the segmented MRI using GLCM stats.

Step 6: Evaluating the training and testing images.

Step 7: RNN classifier is introduced to help classify the state of the segmented brain regions.

Step 8: Compare Ground truth image.

Step 9: Setting the classifier label as 0 and 1 and apply state.

if (result == 1) helpdlg ('tumor is found in brain') disp ('Malignant')

else helpdlg ('No tumor is reported in brain') disp ('Benign')

end if (choice ==3)

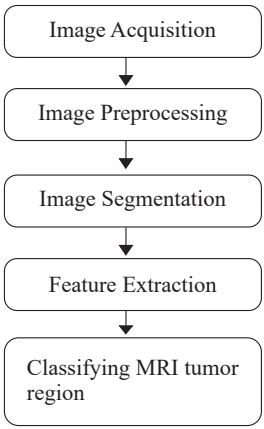
close all

return

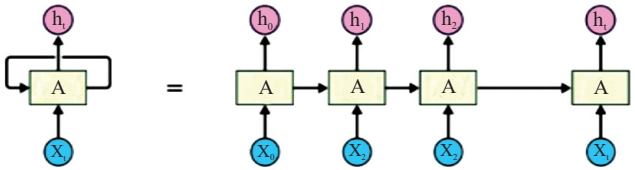
end

Step 10: Return the final results.

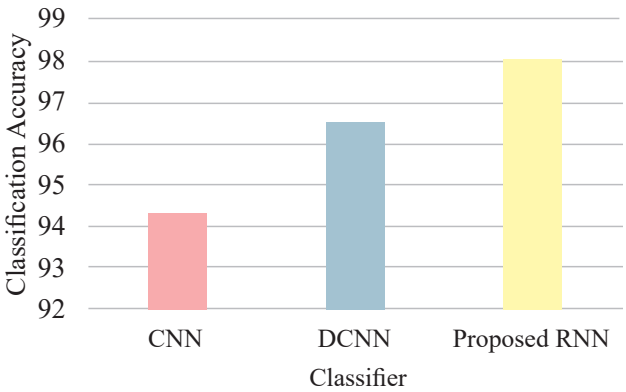
*(***Fig. 10** - *Algorithm)*



*(***Fig. 11** - *Flow of the architecture)*



*(***Fig. 12** - *RNN Architecture)*



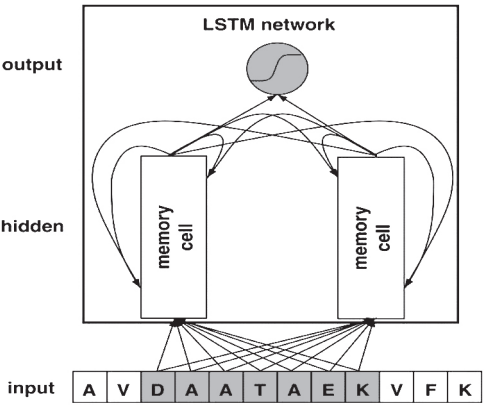
*(***Fig. 13** - *Classification accuracy of the proposed RNN as compared to the existing CNN and DCNN classifier)*

4. "Fast model-based protein homology detection without alignment:"

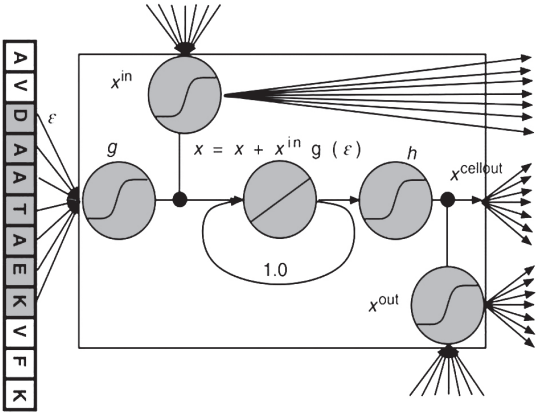
The study comes as an upgrade on methods being used prior to this. The fast analysing of proteins like amino acid with the genome sequencing. Earlier studies were based on Support vector machines which using alignment to compare similarities among sequences.

The proposed algorithm is an RNN architecture which is aimed to overcome shortcomings of alignment-based classifications as they suffer badly when the sequence is shuffled or there is a genetic recombination.

The RNN architecture is poised to be able to extract dependencies between sub sequences and also correlations within them. RNN are also able to extract global sequence characteristics which can be used down the line along with dependencies among amino acids which generally ranges over long intervals in the sequence. [13]



*(***Fig. 14** - *The proposed LSTM architecture, shown with memory cells)*



*(***Fig. 15** - *Memory cells inside the LSTM architecture)*

Method	M	P	V	S	ROC	ROC50	Time
(a) PSI-BLAST	-	-	-	-	0.693	0.264	5.5 S
(b) FPS	-	-	-	-	0.596	-	6800 s
(c) SAM-T98	+	-	-	-	0.674	0.374	200 s
(d) Fisher	-	-	-	+	0.887	0.250	>200 s
(e) Mismatc h	-	-	-	+	0.872	0.400	380 s
(f) Pairwise	-	-	-	+	0.896	0.464	>700 s
(g) SW	-	-	-	+	0.916	0.585	>470 s
(h) LA	-	-	-	+	0.923	0.661	550 h
(i) Oligomer	-	-	-	+	0.919	0.508	2000 s
(j) HMMSTR	-	+	+	+	-	0.640	>500 h
(j) Mismatch-PSSM	-	+	+	+	0.980	0.794	>500 h
(j) SW-PSSM	-	+	+	+	0.982	0.904	>620 h
(k) LSTM	+	-	+	-	0.932	0.652	20 s

*(**Fig. 16** - The first comparison based on SCOP benchmark puts LSTM a fast 2nd only to PSI-BLAST which is a sequencing tool)*

The results gained through LSTM are quite good and outperform other architectures by many folds. Upon further testing with other databases, LSTM moves on top returning the best accuracy and least error. [13]

Method	ROC all	ROC remote	Time
PSI-BLAST	0.80	0.69	50h
SAM 3.5	0.85	0.76	1200h
LSTM	0.88	0.79	27h

*(**Fig. 17** - Results on PFAM database)*

Method	Q	Method	Q
NN	41.8	SVM	45.2
LSTM	51.7		

where 'NN' means neural network and 'SVM' support vector machine. LSTM yields the highest accuracy.

*(**Fig. 18** - Result on Ding and Dubchak dataset for different machine learning methods)*

Method/motif	Sensitivity	Specificity	Balanced Error
PROSITE	85.91 (15.62)	99.94 (0.15)	7.08 (7.79)
LSTM	98.24 (3.55)	99.79 (0.19)	0.99 (1.82)
Motif	86.82 (9.2)	99.93 (0.16)	6.63 (4.59)

*(**Fig. 19** - Result on PROSITE protein classification)*

Conclusion

Going through the above-mentioned case studies and published papers one can easily grasp the supremacy of deep learning networks, Recurrent neural networks here to be precise. The extensive study and results show the possibilities RNN holds in the medical science bubble.

The findings show how close a deep learning model/ architecture is to giving concrete outputs which often at times is a better take and avoids any human error. The ability of RNN to hold previous outputs and learn from them for a more relevant outcome in subsequent layers is commendable, the memory cell may be limited by the length of data it can hold but it still is able to understand the sequences and meaning of the problem or the data input to give commendable results.

To conclude, the growth seems only upwards and more accurate in the coming times, while also decreasing the response times which is very essential where every lost minute, if not second, can be a major deciding factor.

References

[1] https://machinelearningmastery.com/an-introduction-to-recurrent-neural-networks-and-the-math-that-powers-them/

[2] https://www.sciencedirect.com/topics/engineering/recurrent-neural-network

[3] https://machinelearningmastery.com/an-introduction-to-recurrent-neural-networks-and-the-math-that-powers-them/

[4] https://www.bioinf.jku.at/publications/older/2604.pdf

[5] https://www.sciencedirect.com/topics/engineering/recurrent-neural-network

[6] Chapter – 6, Artificial Intelligence in Data Mining Theories and Applications, 2021. D. Binu and B. R. Rajakumar

[7] https://d2l.ai/chapter_recurrent-modern/gru.html

[8] https://clay-atlas.com/us/blog/2021/07/27/gru-en-introduction-note/

[9] https://link.springer.com/article/10.1186/s40537-020-00351-4

[10] Recurrent neural networks in computer-based clinical decision support for laryngopathies: an experimental study - PubMed (nih.gov)

[11] Predicting future medical diagnoses with RNNs using Fast AI API from scratch | by Sparkle Russell-Puleri | Towards Data Science

[12] C7983019320.pdf (ijitee.org)

[13] Fast model-based protein homology detection without alignment | Bioinformatics | Oxford Academic (oup.com)

Math, Optimization, Algorithms behind AI

Chitharanjan Billa

Solutions Architect
(Oracle/Google-Cloud/Big Data/Open-Source/DevSecOps),
22nd Century Technologies Inc., Virginia/Washington D.C.,
District of Columbia, USA
chitharanjan@hotmail.com

***Abstract* --**Artificial Intelligence popularly known as AI is one of the 5 major disruptive technologies in this age [1]. Other four technologies are: *Distributed Ledger Technology, 3D Printing, Virtual Reality and Internet of Things.* Out of these five technologies, AI is advancing very fast across the globe with new industries adapting this technology. Particularly there is lot of emphasis on using AI in self driving (autonomous) cars, with TESLA being prime example, using AI in drug discovery and testing so that new drugs could be developed and delivered in a rapid phase. In addition to these 2 key areas, other areas such as predicting climate change, edge and quantum computing are other 2 essential areas in AI usage.

This white paper depicts the relationship between AI and mathematics, optimizing the algorithms (by applying efficient mathematical techniques). It is vital to appreciate that the more we understand mathematics, AI applications could be built effectively thus solving many critical issues for the better good of humanity. The 4 major areas in mathematics (Calculus, Algebra, Statistics, Probability) which are the key pillars for AI are briefly discussed along with their practical usage in AI. Optimizing algorithms is another area which needs immediate attention else most of the data science and AI projects are set to fail or abandoned in the development phase itself.

Keywords*: AI, Mathematics, Algebra, Calculus, Statistics, Probability, Optimization, Machine Learning, Algorithms,*

Mathematics in Artificial Intelligence:

As discussed earlier, there is a strong correlation between Artificial Intelligence and Mathematics as mathematics is used in Machine Learning algorithms.

The four mathematical areas (pillars) which are extensively used and applied are:

1. Calculus 2. Algebra 3. Statistics 4. Probability.

Calculus*:* Calculus is the mathematical study of continuous change, in the same way that geometry is the study of shape, and algebra is the study of generalizations of arithmetic operations.

It consists of two major areas: Differential Calculus (Differentiation) and Integral Calculus (Integration). Differentiation deals with rates of change (delta), whereas Integration is related to summarization, accumulation of quantities between boundaries. These two branches make

use of the fundamental notions of convergence of infinite sequences and infinite series within the boundaries. Calculus deals with changes in parameters, functions, errors, and approximations. Working knowledge of multi-dimensional calculus is imperative in Artificial Intelligence.

Application of Calculus in Artificial Intelligence: Best example would be calculating and deriving the weights in the artificial neural network, which is also known as back propagation, requires a fair amount of calculus. In addition to this example, there are several calculus applications in Artificial Intelligence technology.

Algebra*:* Artificial Intelligence uses algebra as fundamental method of calculating and interpreting. Algebra includes not only most familiar mathematical operations such as addition, subtraction, multiplication, and division, but also complex mathematical operations such as matrix multiplication, factorial calculations etc.,

Linear algebra: It is one of the important mathematical tools in computational standpoint. This is the foundational mathematical concept in deep learning and machine learning subjects. Calculations based on matrices and vectors is very important in machine learning algorithms which are eventually used in Artificial Intelligence.

Tensors: An important algebraic function is Tensor [2]. Tensors are about an N-Dimensional array (where n>2) of numbers. These Tensors are arranged on a regular grid with N-axes. This algebraic concept is very important in Machine Learning and Deep Learning. Several numerical applications such as finding missing numbers in a data set use Tensors. Tensors provide a natural and summarizing mathematical framework for solving problems and formulating in areas of physics such as relativity, fluid mechanics and elasticity.

Statistics:

Artificial Intelligence (AI) without statistics is unimaginable. Statistics is primarily used for prediction, analysis of huge data sets across the enterprises and government agencies. Government agencies use descriptive statistics to summarize the population data. Some of the common and well used numerical descriptors include *median, mean, and standard deviation* (SD) for flowing data (like income). Frequency distribution is used to categorize data and provide meaningful summaries such as percentages based on demographics etc.,

Statistical inference process is utilized for data analysis to extrapolate date assets of an underlying distribution of probability. This concept is widely used in data modelling, cluster analysis etc.,

Probability:

Probability is solely about occurrence of an unbiased event. Probability concepts have been given an axiomatic mathematical formalization in probability theory. Probability is used in Mathematics, gambling machines, machine learning, artificial intelligence, game theory etc.,

Relation between Artificial Intelligence and Machine Language

Supervised Learning: Supervised learning is a machine learning task where the system takes trained data as input and generates output after performing mathematical calculations. Based on the accuracy of the output, the algorithm gets self-adjusted until an acceptable accuracy is obtained with minimal or no errors. Supervised learning is heavily used in solving and getting meaningful insights in data mining applications.

Supervised learning can be separated into two types of problems when data mining—*classification and regression.* The term "supervised" is used as the data sets, inputs and outputs are known and the data sets are trained for desired outcomes.

Classification is the term used in statistics to categorize input data (test data) based on machine learning algorithm. Eventual outcome of this task would be to conclude how the entities in the datasets could be categorized. Common classification algorithms include Logistic Regression, Naïve Bayes, Decision Tree, Neural Network, Random Forest

Regression: In regression, algorithms try to create or update relationship between dependent and independent variables. This is commonly used to make projections, such as for sales revenue for a given business. Linear regression, logistical regression, and polynomial regression are popular regression algorithms.

Unsupervised learning: In unsupervised learning patterns, labels in the data sets are completely agnostic. In unsupervised learning there are two principal methods which are Neural Probabilistic Methods and Neural Networks. In probabilistic methods there are 2 important types of analysis: Principal Component Analysis and Cluster Analysis. **Principal component analysis:** Enterprises and government agencies are realizing huge data sets in the past decade and this trend would continue to grow. Main problem with huge data sets is that they are very difficult to interpret. Principal component analysis widely used in health industry and financial industry. In medical field it is used in neuroscience where it's used to generate probability of a neuron if there is a probability of generating an action potential [3]. Main principle behind this technique is this mathematical computation reduces dimensionality in data sets and reducing the information loss.

Cluster analysis: Some applications of unsupervised machine learning techniques are:

Clustering algorithms which repeatedly divide large datasets (both structured and unstructured) into discrete groups based on their comparisons and similarities. One major usage is "anomaly detection" which can detect outlier data points in the input dataset. This algorithm could be used in finding fraudulent transactions [4]. After applying this technique to the data set, we can realize clean data as shown below which clearly categorize the data elements.

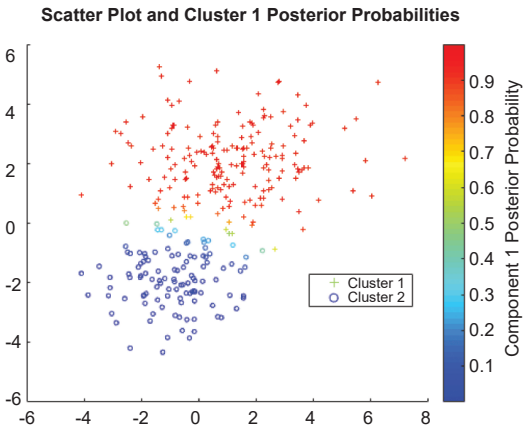


Fig.1 *(Source: mathworks.com)*

Algorithms used in Artificial Intelligence.

There are several algorithms used in Artificial Intelligence technology. Some prominent algorithms are, *Search algorithms* (Google heavily relies on their search algorithms both for public and commercial usage), *Machine learning algorithms* (as discussed above, machine learning is a crucial element in Artificial Intelligence), *Deep learning algorithms* (these algorithms compete with human

reasoning and try to produce similar results especially in comparison standpoint).

Search algorithms are the most primitive and known algorithm which is being used to solve several problems. In solving these problems, algorithms are required to achieve a certain output, which could be a best match in the google search based on the key word, or optimal path when trying to navigate between point-A to point-B. These techniques could be well used in gaming platforms, best outcomes in search platforms and search engines.

Deep learning algorithms: Deep learning algorithms provide further insights about the outcomes. Example – a current algorithm may detect if the object is moving or not, by applying deep learning algorithms to the device, it can further detect it's a human or an animal. Further optimizing this algorithm will produce outcomes such as types of animals, color of the animal etc., Deep learning algorithms are becoming increasingly popular and useful in health sciences, manufacturing industry and other prominent industries where safety of human beings are a concern.

*Machine learning algorithm*s use mathematical and statistical approaches to train data. These ML algorithms are used to create/update relationships in the existing data and provide meaningful insights to the customers and enterprises to make meaningful decisions.

These algorithms are also used to make predictions (trend analysis) with the available data set and forecast the trends. In addition to earlier discussed supervised and unsupervised machine learning algorithms, there is a third type of unsupervised learning which is Reinforcement learning. Reinforcement learning is related to psychology and behavioral patterns and reward system. This algorithm rewards the underlying device (example: a robot which

helps in fitting tires gets rewarded for proper alignment but gets penalized for wrong fit. Over the course of time, the algorithm gets self-optimized so that the robot gets rewarded all the time)

Biology intuitive algorithms – These are the types of algorithms that take inputs from the surrounding world and space and predict estimations and behaviors of the subjects which include both living and non-living objects such as celestial stars. Best example would be "Ant colony optimization" [5] where an artificial ant is a normal computational vehicle which would initiate a search using best optimized path to generate an output (gather food is the outcome with ants).

Applications of AI where Math is used, and Algorithms need continuous optimization:

Automation: Automation is being used from several decades, simple example would be an alarm clock which is set to ring every day at a particular time or many times. Adding AI to the clock alarm would benefit individuals by automatically stopping the alarm during the holidays and weekends or customize to match the ensuing occurrence of events or modifications. These events could be either unanticipated or scheduled.

Machine efficiency: AI can help control a device or a running computer in such a manner as to achieve maximum efficiency both in terms of power utilization and CPU usage. AI bases computer systems such as Grid Infrastructure shuts down the machines which are not in use or be in hibernation mode depending on the usage. Goal is to optimize the power and resources usage per second level. Unused resources are diverted to the areas where the load is needed which intelligently done by the software run in AI algorithms.

Organizational Administration: Be it a big industry or a hospital where several hundreds of thousands of resources work continuously and scheduling becomes a burden thus endangering human lives. AI could be effectively and optimally utilized in these scenarios where needed resources could be deployed where the need is required utmost. Embedding AI in scheduling and personal calendars we could realize immediate benefits.

Protection Systems: Protection systems such as warning bells, smoke alarms use inherent AI to certain degree which we all are used from several decades. These safety techniques are being enhanced to a greater extent and now showing up with intelligent security systems such as alerting the homeowner when a new person tries to enter the premises. These protective benefits were being utilized massively in Cybersecurity area also where petabytes of logs are scanned, categorized, auto reviewed against known vulnerabilities, labelled, and alerted to prevent major cyber-attacks.

Internet of Medical Things: This is another key area where artificial intelligence is embedded in medical devices to proactively monitor the patients and helping them to schedule their visits and if needed provide proactive advice by constantly monitoring their vital parameters. In the ongoing COVID-19 pandemic situation where virus mutations are occurring rapidly, these devices would help sending the latest genomic material for sampling purposes very quickly thus enabling the scientists to assess the mutations and come up with appropriate drugs.

Medical Diagnosis: This is one of the key areas where AI could be used extensively. AI could help doctors in diagnosing a disease or provide additional opinion. Studies show that AI could diagnose little better than the actual radiologist providing better outcomes to the patient with

early detection capabilities. Any missed or wrong diagnosis could result in heavy consequences to the patient and the hospitals involved. The underlying algorithms used in diagnosis should be constantly reviewed and optimized with lots of data sets generated on daily basis. Below chart shows AI superiority over radiologist diagnosis which is a promising outcome for AI utilization [6].

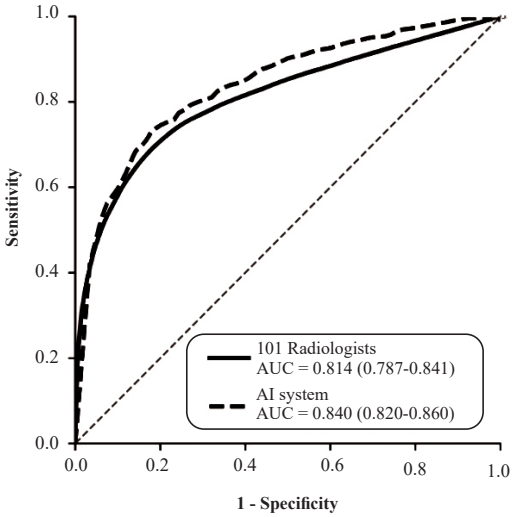


Fig. 2 *Medical Diagnosis*

Monetary Fraud detection: Financial fraud is rampant across the globe with identity theft being primary concern. Imposters could siphon millions of dollars pretending to be the authentic users. Artificial Intelligence is the best solution to prevent, investigate financial fraud starting from common man to business using mathematical algorithms. Cluster analysis, Bayesian networks, link analysis, decision theory are some of the critical algorithms used in monetary fraud detection [7]. Insurance companies are heavily relying on these mathematical algorithms to detect fraudulent claims submitted by both practitioners and patients.

Customer service: The helpdesk or the phone banking helpline is leveraged by AI which mimics human and provide all the essential operational support to the customers by providing basic answers such as current balance, most recent transactions etc.,

CONCLUSION

It's illustrated that there is a strong relationship between Artificial Intelligence and Mathematics and the need for continual optimization of the underlying algorithms is demonstrated.

REFERENCES

[1] Technology HQ (2021). 5 Disruptive Technologies Shaping Our Future.

[2] Web resource. https://www.technologyhq.org/5-disruptive-technologies-shaping-our-future/

[3] Rowland, Todd and Weisstein, Eric W. "Tensor." From MathWorld--A Wolfram

[4] Web Resource. https://mathworld.wolfram.com/Tensor.html

[5] Ian T. Jolliffe and Jorge Cadima (2016). Principal component analysis: a review and recent developments.

[6] Web Resource: https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202

[7] Suresha HP (2021). Cluster analysis in Machine learning Mammography: Comparison With 101 Radiologists.

[8] Web Resource: https://medium.com/mlearning-ai/cluster-analysis6757d6c6acc9

[9] Ant colony optimization algorithms." Wikipedia, Wikimedia Foundation, 11 January 2022,

[10] Web Resource: https://en.wikipedia.org/wiki/Ant_colony_optimization_algorithms

[11] Alejandro Rodriguez-Ruiz, et al., (2019). Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. Web Resource: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6748773/

[12] Data analysis techniques for fraud detection." Wikipedia, Wikimedia Foundation, 08 December2021, https://en.wikipedia.org/wiki/Data_analysis_techniques_for_fraud_detection

Quantum integration to AI for higher fidelity in medical research and deployment

Sarthak Chakraborty

Department of computer Science and engineering
CMR University (CMRU)
Bangalore, India
sarthak.chakraborty.144@gmail.com

Abstract --Since the dawn of Artificial Intelligence (AI) in medical science for achieving application in curing cancer cells to making personalized medicine. Medical sector has wide application of AI as per today's team of research and development, organizations and big technology giants. Researcher are widely accepting it as the next revolution in medical history where **precision is the key to the cure**. As in 2022 today we can see that none of the systems are yet ready to deploy on its own without the human intervention. It's still in research phase due to many factors of collision to the completion. Researchers starts to hit the saturation point of the application, yet the error being a big leap to be justified for use in daily life. In Machine Learning we have

seen that user or developer of the code has to retrospect on the data, and that is necessary to term it down and why and where a change has to be made.

But in AI the major problem of error detection and fixing is, its algorithm is a big chunk of branches from its own learning where a developer or a group of researcher needs to deploy a lot of man power and resources to handle this anomaly and that's not it, after finding one error in the due course of time, like in a sample set of medical records and their processing, this error mounts to millions and billions, where branches get off the chart and system get faulted. These systems face another saturation point of computation errors, which cannot be sufficed by classical hardware so here comes a whole a set of integration known as Quantum AI or (Q-AI).

This gives ten million times more processing power to the AI to process, because of its, Qubit processing. Quantum Computation derives its routes from the fundamentals of quantum mechanics and distribution probability theory where it's not only fast but highly intuitive, according today's research from one of the tech giants such as Google and IBM are set to deploy more qubits which will directly affect the development of the cloud ecosystem of data processing. Along with it integrating the power of AI will make it the ultimate tool for medical history in the medical environs. Medical ecosystem is such an arena where bigger the data pool processing more the accuracy and benefits of the developed product or applications.

Medical science is an extremely delicate sector and we are talking about people and life of the masses. So, this is where the concern of efficiency of AI comes in and we can develop it. Quantum Artificial Intelligence is the answer. For a researcher while developing fidelity will be just a number but when deployed it will be a mass, we are

talking about who will be at the risk of fault of efficiency. In medical research the bonds of the elements for cure through medicine can be completely simulated where it will show risk analysis to cure rate. Most importantly it will benefit two major sectors that is personalized medicine and research of cells.

A Systematic Review on Role of Machine Learning Techniques Prediction of Covide-19 Cases

Lalita Charpe

M.Tech Scholar, Department of Computer Science and Engineering, Tulshiram Gaikwad College of Engineering, Nagpur, Maharashtra, India

Jayant Adhikari

Assistant Professor, Department of Computer Science and Engineering, Tulshiram Gaikwad College of Engineering, Nagpur, Maharashtra, India

Abstract

SARC-CoV-2, the novel corona virus that is responsible for COVID-19, has caused havoc around the world, with patients presenting with a compressive range of issues that have prompted health-care professional to investigate innovative technology solutions and treatment planned. Some of the organizations have been quick to embrace and redesign Artificial Intelligence (AI) based technologies in response to the challenges affected by the COVID-19

disease. Artificial Intelligence (AI) based technologies have played a significant role in solving complicated problems, and several organizations have been quick to acquire and redesign these mechanism. A structure evaluation of the literature on the role of artificial intelligence (AI) as comprehensive and decisive technology in the fight against the COVID-19 problem in the fields of hygienic, diagnosis, and ailment progression was the main goal of this examination.

In accordance with PRISMA (Preferred Reporting Items For Systematic Reviews and Meta-Analysis) guidelines, a methodical search of the PubMed, Web of Science, and CINAHL datasets was managed between December 1, 2019 and June 27, 2020 to recognize all imaginably relevant studies that were published and made easily available online between December 1, 2019 and June 27, 2019. The search syntax was created by incorporating terms that were specific to COVID-19 and AI. As part of this systematic review, we gathered papers from the current COVID-19 literature that made use of artificial intelligence-based strategies to provide discernment into various COVID-19 topics. Our findings point to relevant components, data types, and COVID-19 resources that can be used to facilitate clinical and translational research and can be used to utilize detail information for future study.

Keywords: *COVID-19, coronavirus, SARCCoV-2, artificial intelligence, machine learning, deep learning, systematic review, epidemiology, pandemic, neural network.*

Deep Learning Solutions for Skin Melanoma Detection

Ashish Anil. Meshram

Tulsiramji Gaikwad–Patil College of Engineering & Technology Nagpur, India
ashishameshram1@gmail.com

Anup Gade, Abhimanyu Dutonde

Tulsiramji Gaikwad–Patil College of Engineering & Technology Nagpur, India
deanacademics@tgpcet.com

Abstract: Skin cancers, a concerning public fitness quandary, with over 5,000,000 newly diagnosed cases each 12 months, simply inside the United States. Generally, skin cancer is of two sorts: melanoma and non- melanoma. Melanoma also called as Malignant Melanoma is the nineteenth maximum regularly happen most cancers in men and women. It is the deadliest form of skin cancer [1]. In the year 2015, the worldwide prevalence of melanoma changed into approximated to be over 350,000 instances, with around 60,000 deaths. The most widely wide-spread

non-melanoma tumours are squamous mobile carcinoma and basal mobile carcinoma. Non-melanoma pores and skin most cancers is the fifth most often occurring most cancers, with over 1 million diagnoses global in 2018 [2]. As of 2019, extra than 1.7Million new instances are predicted to be identified [3]. Even although the mortality is significantly high, however while detected early, survival rate exceeds 95%. This motivates us to give you a solution to store millions of lives through early detection of pores and skin most cancers. Convolution Neural Network (CNN) or ConvNet, are a category of deep neural networks, essentially generalized version of multi-layer perceptions. CNNs have given Maximum accuracy in visible imaging duties [4]. This challenge targets to expand a skin cancer detection CNN version that can classify the pores and skin most cancers kinds and assist in early detection [5].

The CNN type version might be developed in Python the use of Keras and Tensor flow within the backend. The model is advanced and tested with one-of-a-kind community architectures through varying the kind of layers used to teach the network such as however now not restricted to Convolution layers, Dropout layers, pooling layers and dense layers. The version may also employ Transfer Learning strategies for early convergence. The version will be tested and educated on the dataset amassed from the International Skin Imaging Collaboration (ISIC) challenge archives.

Keywords: *Neural community, Skin most cancers, Deep mastering, Machine studying, Cancer detection, Cancer analysis Convolution neural community CNN, Melanoma*

EDITOR

Dr. Shobana Padmanabhan

Dr. Shobana Padmanabhan received her Ph.D. and M.S. in Computer Science and Engineering from Washington University in St. Louis, USA. Her undergraduate degree is from BITS, Pilani, Rajasthan. She has over 20 years of experience in research, teaching, and industry. Dr. Shobana's research interests include deep learning, machine learning and domain-specific optimization. Her teaching interests include student-centric active learning. Dr. Shobana has publications in over 20 international conferences and journals, one international patent, one Indian copyright and a book chapter to her credit.

CO-EDITOR

Dr. Vipin Prasad

Dr. Vipin Prasad possesses a master's degree in Physics from Cochin University of Science and Technology with a specialization in Solid State Physics, with optional subjects Solid State Electronics, Thin Film Physics and Laser Physics. Dr. Vipin obtained his doctoral degree from Banaras Hindu University, Varanasi in Laser Spectroscopy (Laser Thermal Lensing).

He has 26 years of Industrial Experience at Department of Defence Production, Ministry of Defence, Government of India. Joined Alliance University as a Professor in Department of Physics.

TECHNICAL COMMITTEE

Dr. Chetan J. Shelke

Dr. Chetan J. Shelke holds a Master's degree (M.E) in Information Technology from PRMIT, Badnera, SGBAU Amravati University. After a two year stint in the college, Dr. Chetan joined the P.R.Pote (Patil) Group of Educational Institutions and worked as head of Dept. (Information Technology) for 10 years, and obtained his Ph.D. from Sant Gadge Baba Amravati University, Amravati. His areas of interest include data science and cyber security.

He has presented and published more than 40 papers in national and international conferences/ journals and has six patents. He has awards from Science, Technology and Engineering Systems Journal and membership at Asia Society of Researchers and a Merit Certificate from Amravati University.

Dr. M.Senbagavalli

Dr. M.Senbagavalli has a doctoral degree from Anna University. She also holds an M.E. in Computer Science and Engineering from Anna University and a B.E. in Information Technology degree from the Periyar University. Her areas of teaching and research include: Data Mining, Big Data Analytics, IoT, Computer Networks, Object Oriented Programming (C++ and Java), Object Oriented Analysis and Design, Visual Programming, C# and .Net, Software Testing, Software Engineering and Quality Assurance, Design and Analysis of Algorithms. She has presented and published papers in national and international conferences/ journals focusing on Data Mining, Big Data Analytics, Cloud Computing and Opinion Mining.

About the conference

The one-day symposium on Artificial Intelligence (AI) for Medical Science aims to bring together researchers, practitioners and professionals from academia, industry, hospitals, clinics, and service centers to participate in thoughtful and informative presentations and discussions. On this platform, experts can discuss recent advances in the diagnosis, interventions and the impact of cutting-edge developments in the research and evidence-based practices.

ISBN 978-81-955963-2-4

9 788195 596324

Published by

Center for Social Impact and Research
Alliance University, Bengaluru

ALLIANCE UNIVERSITY
Private University established in Karnataka State by Act No.34 of year 2010
Recognized by the University Grants Commission (UGC), New Delhi

Chikkahagade Cross, Chandapura - Anekal Main Road, Anekal, Bengaluru – 562 106, Karnataka, India.