ABSTRACT

Deep learning provides the healthcare industry with the ability to analyze data at exceptional speeds without compromising on accuracy. These techniques are applicable to healthcare domain for accurate and timely prediction. Convolutional neural network is a class of deep learning methods which has become dominant in various computer vision tasks and is attracting interest across a variety of domains, including radiology. Deep Learning and Machine Learning also plays most motivating space of exploration that become generally well known in wellbeing association. It likewise has a crucial impact to reveal new examples in therapeutic science and administrations affiliation which subsequently obliging for every one of the gatherings related with this field. This undertaking expect to frame a symptomatic model of the normal sicknesses dependent on the manifestations by utilizing information mining method like arrangement in wellbeing space. In this project, we will utilize AI calculations and profound realizing which can be used for health care diagnosis. In this project, we are proposing the disease classification using symptoms and images: Chest x-ray images of lung cancer, covid and pneumonia disease. Accordingly we are recommending the precautions and hospital to the patient.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Human is master in getting data, while machine is master at communicating and handling information. In this project, we propose a model for patient side effect closeness examination by exploiting the machine's capacity to process information. The model utilized patient's portrayals of indications to remove key data and accomplish early expectation and mediation. Consequently, the precision of likeness examination model to a great extent decides the adequacy of infection expectation. Accurately predicting diseases plays a significant role in public health, especially at the early stage which allows patients to take prevention treatments in time. With the growing volume and availability of electronic health records (EHRs), predictive modeling tasks for disease progression and analysis have obtained increasing interest from researchers. The EHR data are temporally sequenced by patient visits with each visit represented as a set of high dimensional clinical events. Mining EHRs is especially challenging compared to standard data mining tasks, due to its noisy, irregular and heterogeneous nature. Recently, deep learning and machine learning approaches have been widely adopted and rapidly developed in patient representation learning.

1.2 Motivation

- Reducing death rate by wrong diagnosis using giving accurate diagnosis.
- To provide diagnosis system this helps to doctors.
- Motive behind proposed work is to achieve higher accuracy over existing work by using machine learning.
- The desire to provide a better and accurate diagnosis.

1.3 Objectives

- Main goal is to build up a framework that helps the clinical specialists to cross confirm their analyzed results of predicted covid19, Pneumonia and lung cancer.
- Another major objective is to make an automated tool based on image processing and ML, which minimizes human effort in predicting the presence of covid19, Pneumonia and lung cancer from image as well as through symptoms.

1.4 Expected Outcome

- The diseases will get classified into four categories, provide precautionary measures and hospital address according to the users location.

1.5 Social Relevance: Clinical decision making

By using ML in predicting the effectiveness of a therapy and optimizing the combination of different therapies will pave the way for personalized treatment.

1.6 Organization of Report

1. The second chapter represents the literature survey.

2. The third chapter includes the proposed system, designing project and block diagram

3. The fourth chapter represents simulation and testing

4. The fifth chapter tells the conclusion and references.

JUSTIFICATION FOR THE PROJECT

This project will focus on classifying diseases based on X-Ray images using deep learning. This project will present an alternate option of classifying the lung diseases apart from the traditional option and assist the medical and healthcare system in doing their job accurately. The main justification behind this project is to increase the accuracy over existing work using deep learning and machine learning models in predicting lung diseases.

CHAPTER 2

LITERATURE REVIEW

1. Title: Symptoms based Early Clinical Diagnosis of COVID-19 Cases using Hybrid and Ensemble Machine Learning Techniques

Details: C Kaushik (2021) IEEE Xplore

Methods: This paper analyses the possibilities of improvement in classification results through ensemble and hybrid approaches. It is observed from the results that K-mode clustering followed by classification based hybrid modeling.
Random Forest, XG Booster

Accuracy: 87.17%

2. Title: Pneumonia Detection from Chest X-ray using Transfer Learning

Details: Sharvari Kalgutkar(2021) IEEE Xplore

Methods: The study highlights Convolutional Neural Networks, develops and trains models such as VGG16, ResNet-50, and InceptionV3, to detect pneumonia with improved testing accuracies.

Accuracy: 93.5%

3. Title: CNN based Covid-aid: Covid 19 Detection using Chest X-ray

Details: Shrinial Singh et al.(2021) Conference IEEE

Methods: This study aims to carry out the task of detecting the disease through radiography images of the human chest. Our deep learning model works on a publicly available dataset and uses the concepts of convolutional neural networks.

Accuracy: 87%

4. Title: Pneumonia and COVID-19 Detection using Convolutional Neural Networks

Deatils: Sammy Militane et al. (2020) IEEE

Methods: The trained VGG-16 model the researchers proposed in this research study for the COVID-19 detection and pneumonia detection on chest x-ray images using the CNN method.

Accuracy : 95%

5. Title: Concatenated and Connected Random Forests with Multiscale Patch Driven Active Contour Model for Automated Brain Tumor Segmentation of MR Images.

Details: Chao Ma, Gongning Luo, and Kuanquan Wang

Methods: employ a feature representations learning strategy to effectively explore both local and contextual information from multimodal images for tissue segmentation by using modality specific random forests as the feature learning kernels. Different levels of the structural information is subsequently integrated into concatenated and connected random forests. multiscale patch driven active contour (mpAC) model is exploited to refine the inferred structure by taking advantage of sparse representation techniques.

Accuracy : 90%

6. Title: Supervised Discriminative Sparse PCA for Com-Characteristic Gene Selection and Tumor Classification on Multiview Biological Data

Details: Chun-Mei Feng, Yong Xu , Senior Member, IEEE, Jin-Xing Liu , Member, IEEE, Ying-Lian Gao, and Chun-Hou Zheng , Member, IEEE

Methods: Principal component analysis (PCA) has been used to study the pathogenesis of diseases.
This paper developed a new PCA method, which is named the supervised discriminative sparse PCA (SDSPCA).

7. Title: A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans

Details: Onur Ozdemir, Member, IEEE, Rebecca L. Russell, and Andrew A. Berlin, Member, IEEE.

Methods: Based entirely on 3D convolutional neural networks and achieves state-of-the-art performance for both lung nodule detection and malignancy classification tasks on the publicly available LUNA16 and Kaggle Data Science Bowl challenges.

Accuracy: 87.4%

8. Title: Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey

Details: Khan Muhammad , Member, IEEE, Salman Khan , Student Member, IEEE, Javier Del Ser , Senior Member, IEEE, and Victor Hugo C. de Albuquerque , Senior Member, IEEE

Methods: survey covers the main steps of deep learning-based BTC methods, including preprocessing, features extraction, and classification, along with their achievements and limitations. Also investigate the state of-the-art convolutional neural network models for BTC by performing extensive experiments using transfer learning with and without data augmentation.

Accuracy:89%

9. Title: Machine Learning Applied in SARS-CoV-2 COVID 19 Screening using Clinical Analysis Parameters

Details: R. F. A. P. Oliveira, C. J. A. Bastos-Filho, A. C. A. M. V. F. Medeiros, P. Buarque, D. L. Freire

Methods: Machine learning techniques, such as Random Forest, Multi-Layer Perceptron, and Support Vector Machines Regression, enable the creation of disease prediction models and artificial intelligence techniques to analyze clinical parameters. Thus evaluated the existing correlations between laboratory parameters and the result of the COVID-19 test, and developed two classification models: the first classifies the test results for patients with suspected COVID-19, and the second classifies the hospitalization units of patients with COVID-19, both according to the laboratory parameters.

Accuracy:96%

10. Title: Prediction of Pneumonia Using Big Data, Deep Learning and Machine Learning TechniquesPrediction of Pneumonia Using Big Data, Deep Learning and Machine Learning Techniques

Details: Swetha K R IEEE XPLORE(2021)

Methods: CNN (Convolutional Neural Networks) stand tall and high in this prediction along with other classifiers. Also, pre-training the CNN models for very large datasets that is for big data of healthcare units stands a high chance for accurate classification. A CNN model which is pre-trained along with an efficient feature extraction technique and various classifiers to classify the positive from negative is considered to give highly accurate results.

Accuracy : 94.3%

CHAPTER 3

DESIGN OF THE PROPOSED SYSTEM

3.1 Introduction
This chapter explores the design of proposed system and its working

In a proposed system, we are detecting disease based on symptoms, reports. Also we are suggesting hospital to patient which is near from its distance.

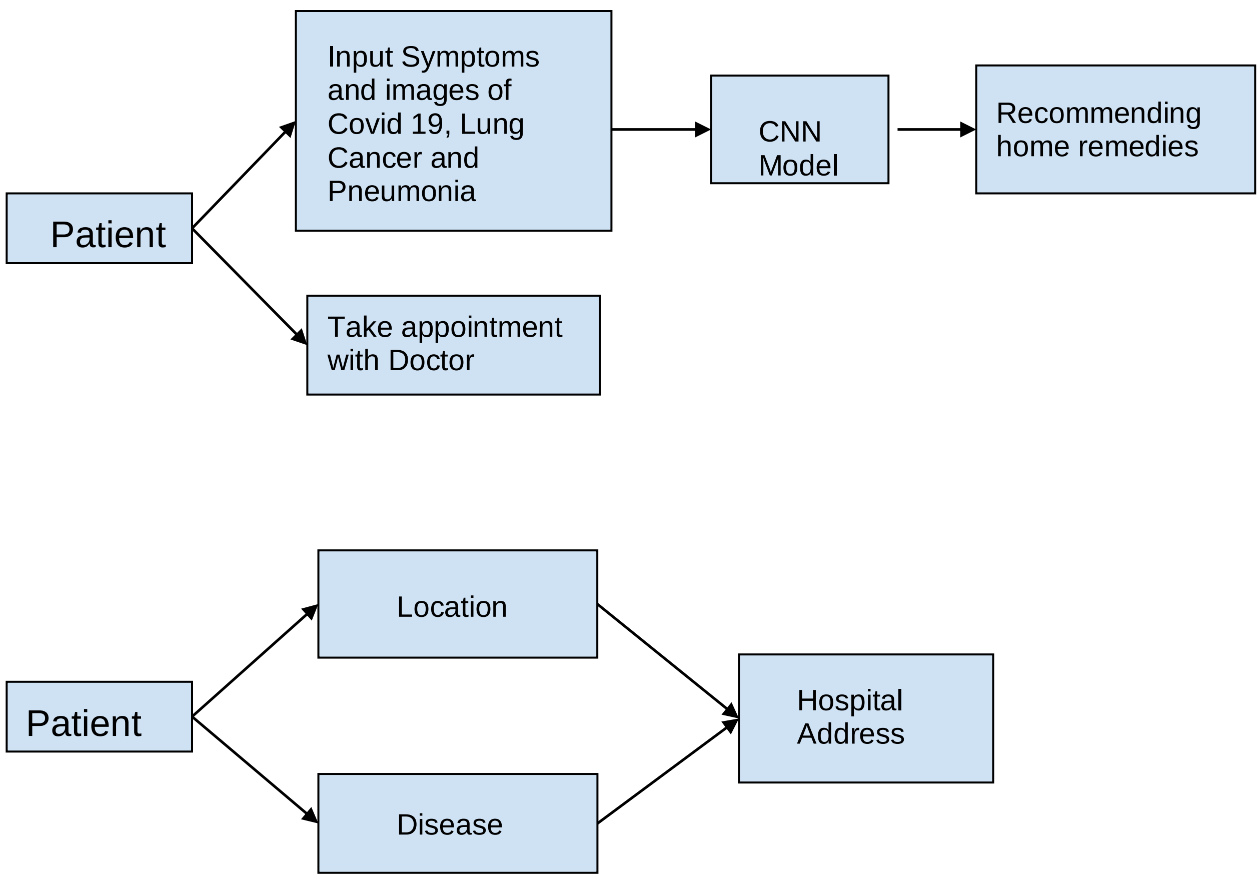


Fig.3.1 Block Diagram of Proposed System

Explanation of the Block Diagram: The patient will give input of symptoms and images of X-ray of the three diseases namely: Covid-19, Pneumonia and Lung Cancer. Using the trained CNN Model, we classify the diseases and recommend home remedies and precautions. The patient will also be able to take appointments with the doctor.

The patient will also provide the input in form of location and disease. Using KNN model, we will provide the patient will nearest hospital address/location.

Algorithm Used:

1. **CNN**

Convolutional Neural Networks (which are additionally called CNN/ConvNets) are a kind of Artificial Neural Networks that are known to be tremendously strong in the field of distinguishing proof just as picture order.

Four main operations in the Convolutional Neural Networks are shown as follows:

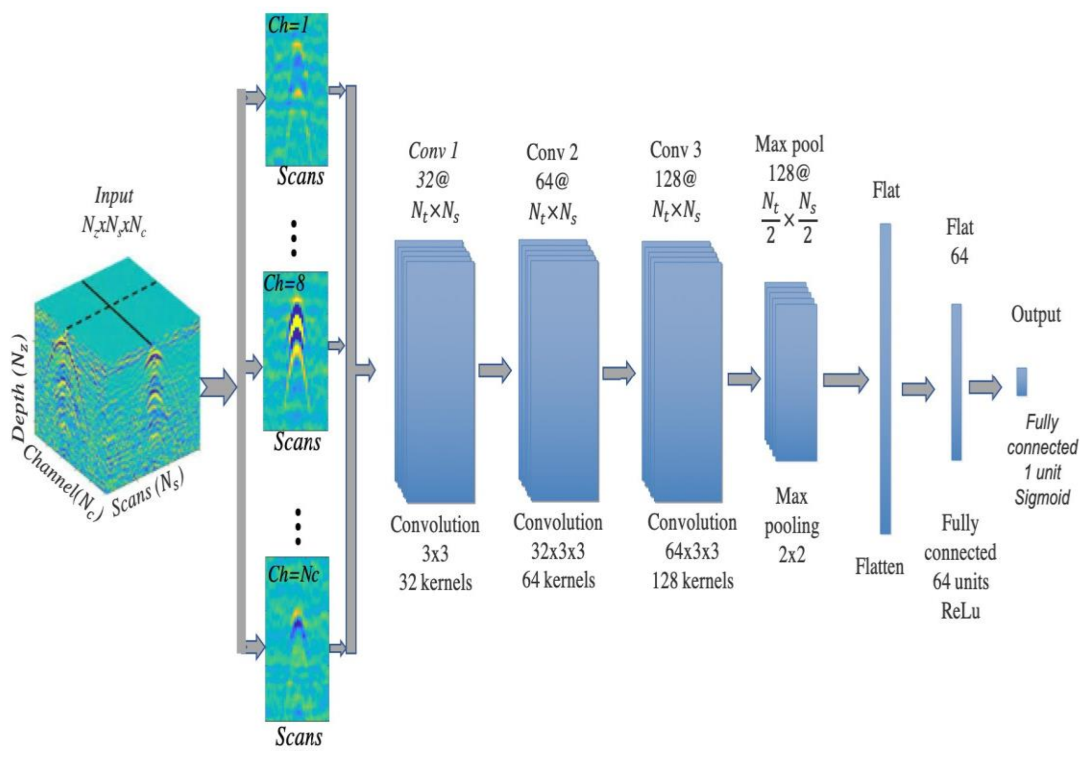


Figure 3.2 CNN Architecture

(i) Convolution

The principle utilization of the Convolution activity if there should be an occurrence of a CNN is to recognize fitting highlights from the picture which goes about as a contribution to the primary layer. Convolution keeps up the spatial interrelation of the pixels This is finished by fulfillment of picture highlights utilizing miniscule squares of the picture. Convolution equation. E very picture is seen as a network of pixels, each having its own worth. Pixel is the littlest unit in this picture grid. Allow us to take a 5 by 5(5*5) framework whose qualities are just in twofold (for example 0 or 1), for better agreement. It is

to be noticed that pictures are by and large RGB with upsides of the pixels going from 0 - 255 i.e 256 pixels.

(ii). ReLU

ReLU follows up on a rudimentary level. All in all, it is an activity which is applied per pixel and overrides every one of the non-positive upsides of every pixel in the component map by nothing.

(iii). Pooling or sub-sampling

Spatial Pooling which is likewise called subsampling or downsampling helps in lessening the elements of each element map yet even at the same time, holds the most important data of the guide. Subsequent to pooling is done, in the long run our 3D element map is changed over to one dimensional component vector.

(iv) Fully Connected layer

The yield from the convolution and pooling activities gives noticeable highlights which are removed from the picture. These highlights are then used by Fully Connected layer for consigning the info picture into various classes predicated on the preparation dataset.

Random Forest

A random forest is a machine learning technique that’s used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

A random forest algorithm consists of many decision trees. The ‘forest’ generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.

The (random forest) algorithm establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome.

KNN

A refinement of the k-NN characterization calculation is to gauge the commitment of every one of the k neighbors as indicated by their distance to the inquiry point, giving more prominent load to nearer neighbors. The KNN classifier recommending the emergency clinic subtleties for infection dependent on the closest distance.

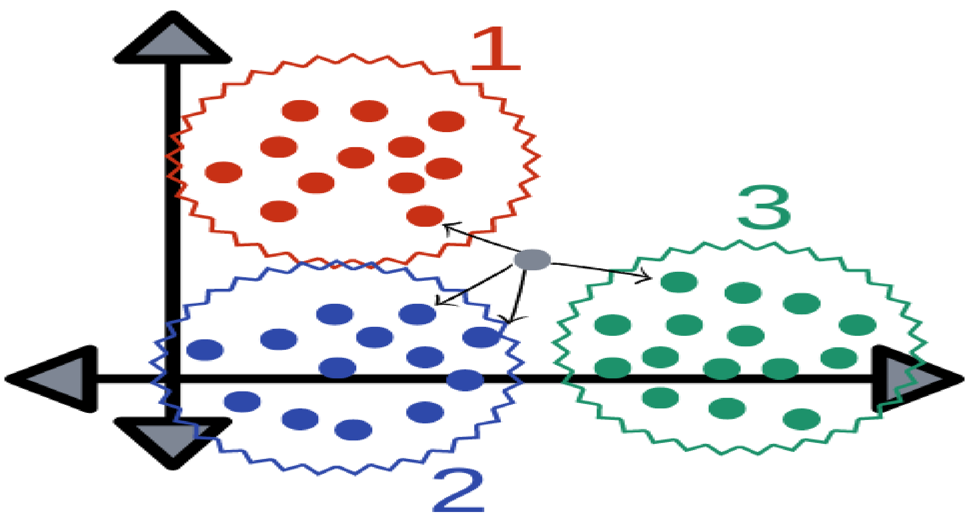


Figure 3.3 KNN Architecture

$d=\sqrt{((x2-x1)^2+(y2-y1)^2)}$

d=distance

x1, x2, y1, y2 = data points.

3.2 Data Flow Diagrams

A **data flow diagram** (**DFD**) is a graphical representation of the "flow" of data through an information system, modeling its *process* aspects. A DFD is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. DFDs can also be used for the visualization of data processing (structured design).

A DFD shows what kind of information will be input to and output from the system, how the data will advance through the system, and where the data will be stored. It does not show information about process timing or whether processes will operate in sequence or in parallel, unlike a traditional structured flowchart which focuses on control flow, or a UML activity workflow diagram, which presents both control and data, flows as a unified model.

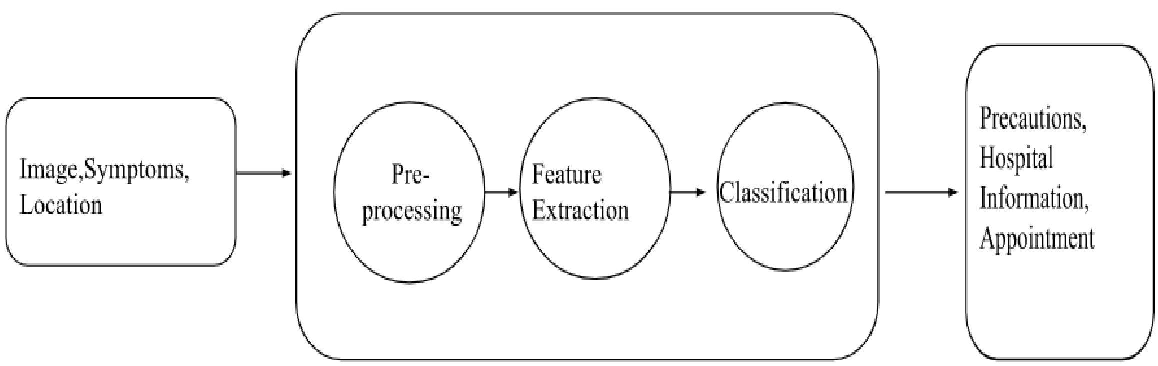


Fig 3.4 DFD

3.3 User Activity Diagram:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams are intended to model both computational and organizational processes (i.e. workflows). Activity diagrams show the overall flow of control. Activity diagrams are constructed from a limited number of shapes, connected with arrows. The most important shape types:

- Rounded rectangles represent actions;
- Diamonds represent decisions;
- Bars represent the start (split) or end (join) of concurrent activities;
- A black circle represents the start (initial state) of the workflow;
- An encircled black circle represents the end (final state).

Arrows run from the start towards the end and represent the order in which activities happen. Hence they can be regarded as a form of flowchart. Typical flowchart techniques lack constructs for ex-pressing concurrency. However, the join and split symbols in activity diagrams only resolve this for simple cases; the meaning of the model is not clear when they are arbitrarily combined with decisions or loops.

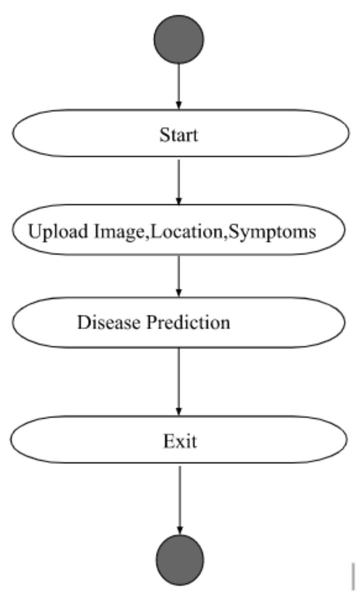


Fig 3.5 User Activity Diagram.

CHAPTER 4

SIMULATION AND TESTING

4.1 Overview of Project Modules

We propose a new Convolutional neural network based disease classification and precautions recommendation model with higher accuracy. We are going to solve accuracy issue in classification of disease images with accurate stage predictions.

4.2 Hardware Testing

Operating System Used: Windows 11 i5 11th generation

RAM (Random Access Memory): 16gb

4.3 Software Testing

Programming Language Used: Python
IDE Used: Spyder

Libraries Used:
1]Tensorflow

2]Keras

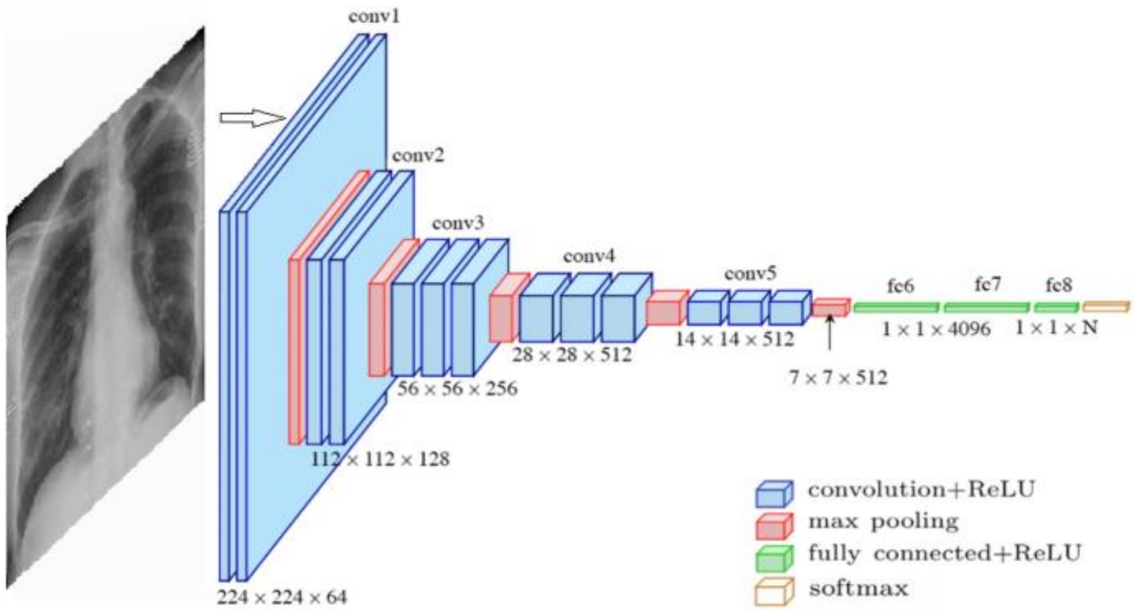
3]Numpy

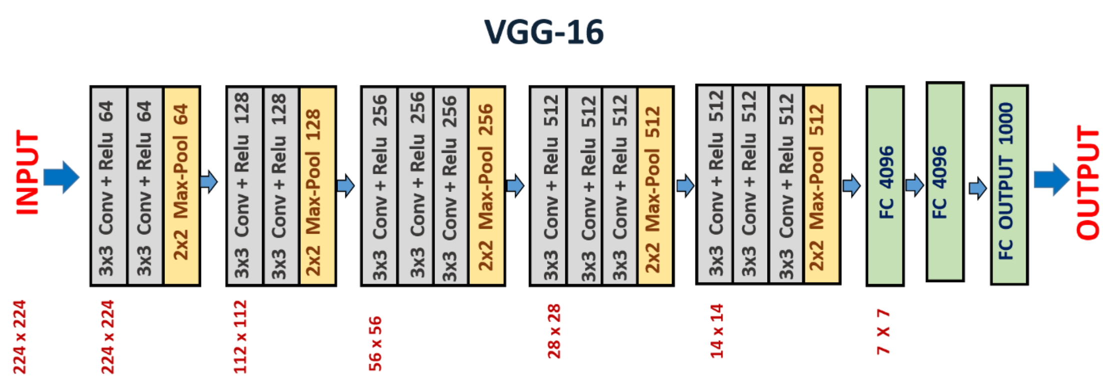
4]Matplotlib

Pretrained Model Used: VGG-16

VGG-16 is a convolutional neural network that is 16 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

Fig 4.1Architecture of VGG-16





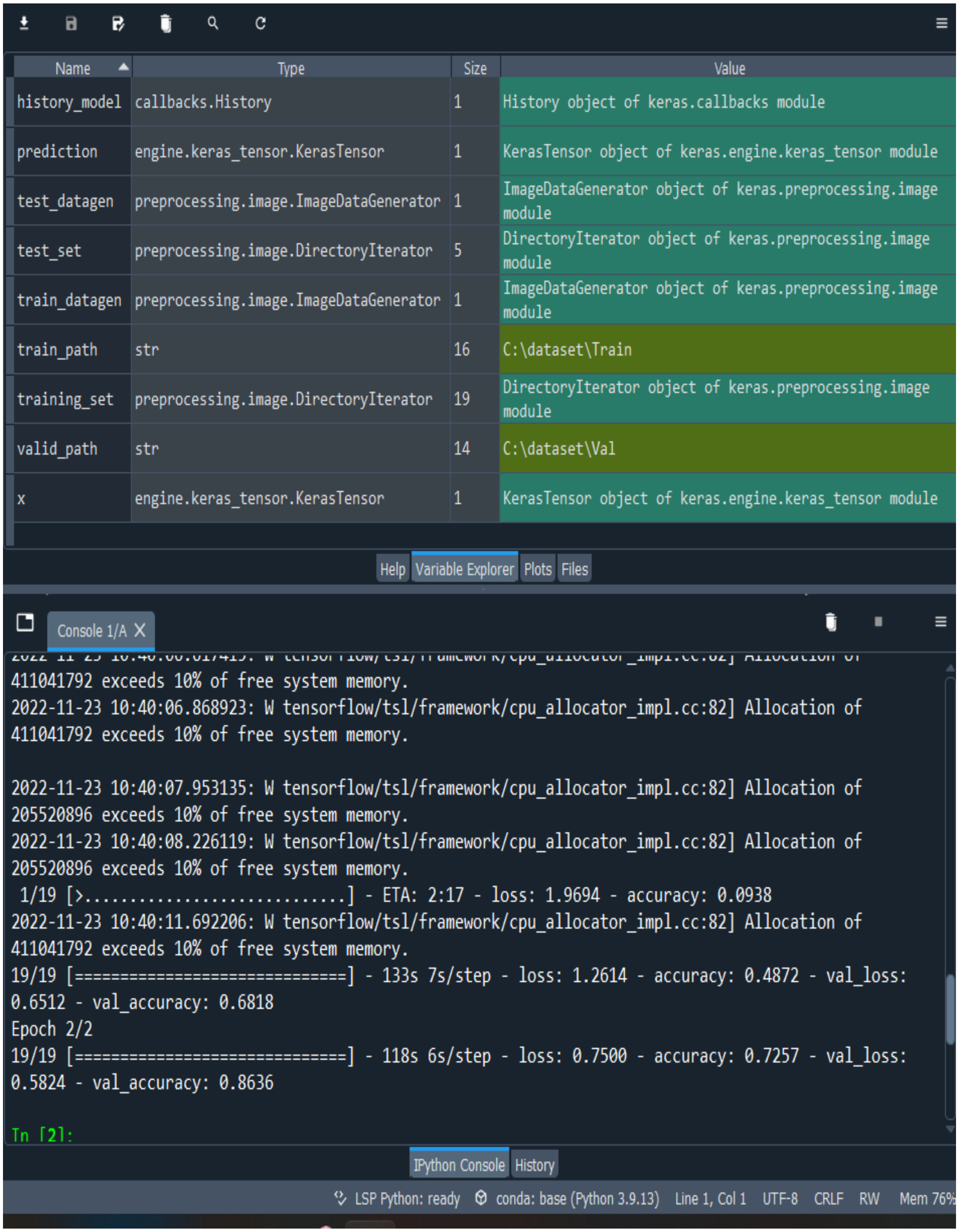


Fig 4.2 Output

CHAPTER 5
RESULT ANALYSIS AND CONCLUSION

5.1 Result Analysis

By using VGG-16 Pretrained model, we have observed that the accuracy by using the dataset is

Cycle	Accuracy
Epoch 1	68.18
Epoch 2	86.36

5.2 Conclusion

We are going to create an accurately predicting classification of the diseases using machine learning and deep learning. The current implementations are going to deal with various techniques with higher accuracy over existing work done.

REFERENCES

[1] C. Ma, G. Luo, and K. Wang, “Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of MR images,” IEEE Trans. Med. Imag., vol. 37, no. 8, pp. 1943–1954, Aug. 2018.

[2] C.-M. Feng, Y. Xu, J.-X. Liu, Y.-L. Gao, and C.-H. Zheng, “Supervised discriminative sparse PCA for com-characteristic gene selection and tumor classification on multiview biological data,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 10, pp. 2926–2937, Oct. 2019.

[3] Onur Ozdemir, Rebecca L. Russell “A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans” IEEE TRANSACTIONS ON MEDICAL IMAGING *last revised 21 Jan 2020.*

[4] Anum Masood, Bin Sheng, Po Yang, Ping Li “Automated Decision Support System for Lung Cancer Detection and Classification via Enhanced RFCN with Multilayer Fusion RPN” IEEE Transactions on Industrial Informatics (Volume: 16, Issue: 12, Dec. 2020).

[5] Khan Muhammad; Salman Khan; Javier Del Ser; Victor Hugo C. de Albuquerque. “Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey” IEEE Transactions on Neural Networks and Learning Systems (Volume: 32, Issue: 2, Feb. 2021).

[6] David N. Louis, Arie Perry, et al. , “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary” , Acta Neuropathol , Springer may 2016

[7] Pär Salander, A Tommy Bergenheim, Katarina Hamberg, Roger Henriksson, Pathways from symptoms to medical care: a descriptive study of symptom development and obstacles to early diagnosis in brain tumor patients, Family Practice, Volume 16, Issue 2, April 1999, Pages 143–148,

[8] McKinney PA,”Brain tumors: incidence, survival, and etiology”,Journal of Neurology, Neurosurgery & Psychiatry 2004;75:ii12-ii17.

[9] V. P. G. p Rathi and S. Palani, “Brain tumor MRI images classification with feature selection and extraction using linear discrimination analysis,”2012, *arXiv:1208.2128.*

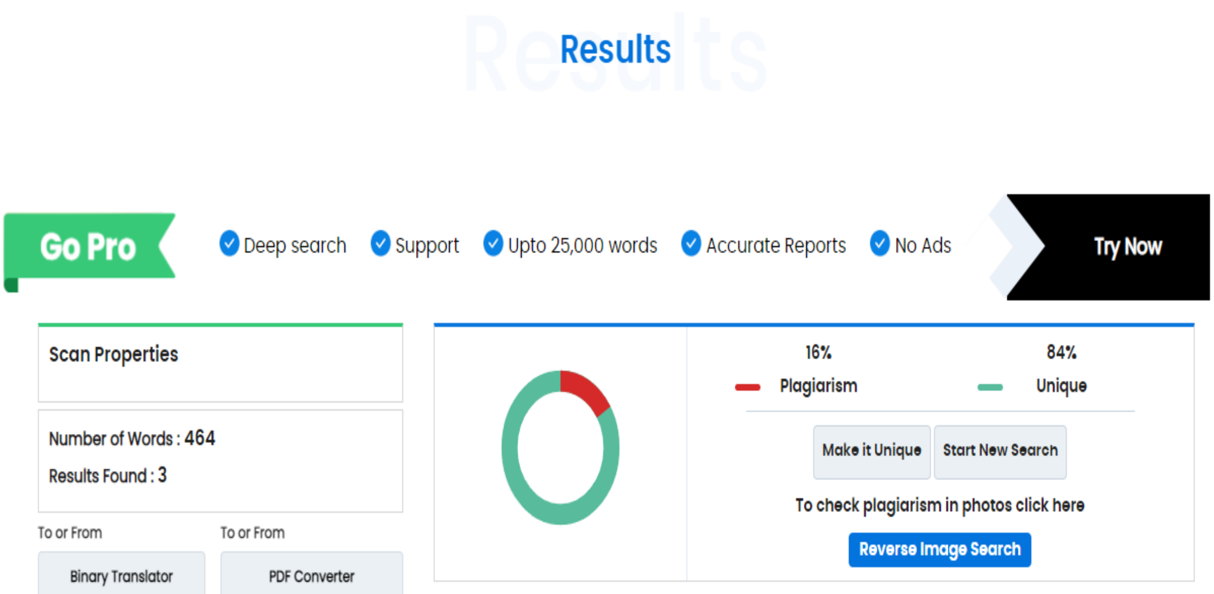
[10] R.F.A.P Oliveria, C.J.A.A. Bastos-Filho, A.C.A.M.V.F Medeiros, P.Buarque “Machine Learning Applied in SARS-CoV-2 COVID-19 Screening using clinical analysis”, IEEE Trans, Vol.19, no. 6 June 2021

APPENDIX

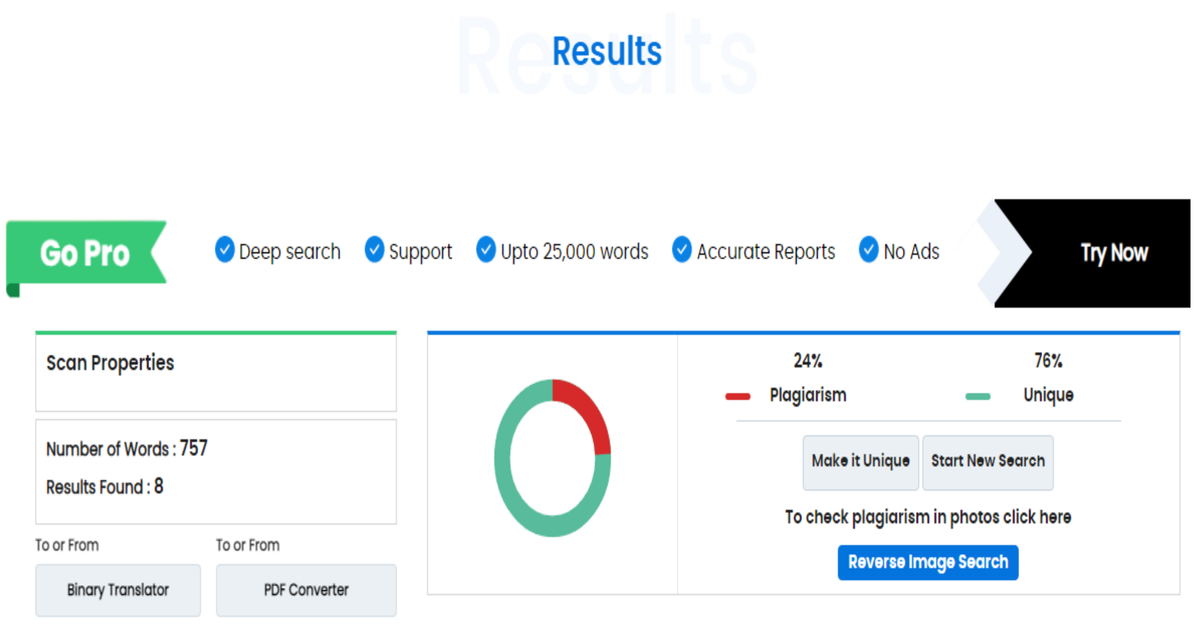
1. Plagiarism Reports

PLAGIARISM REPORTS

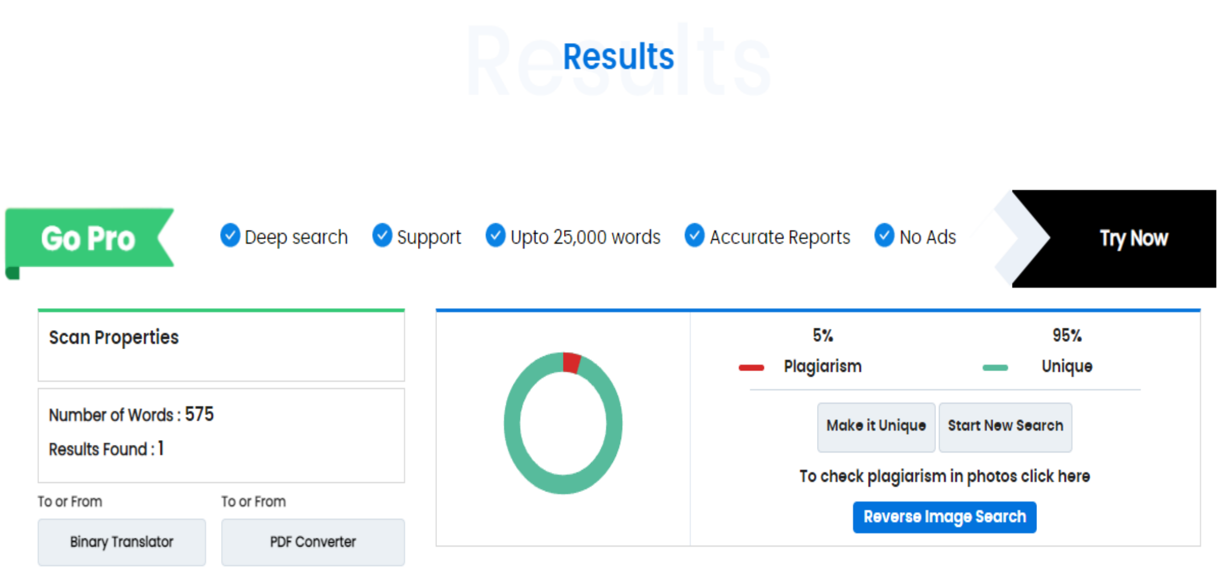
CHAPTER 1:



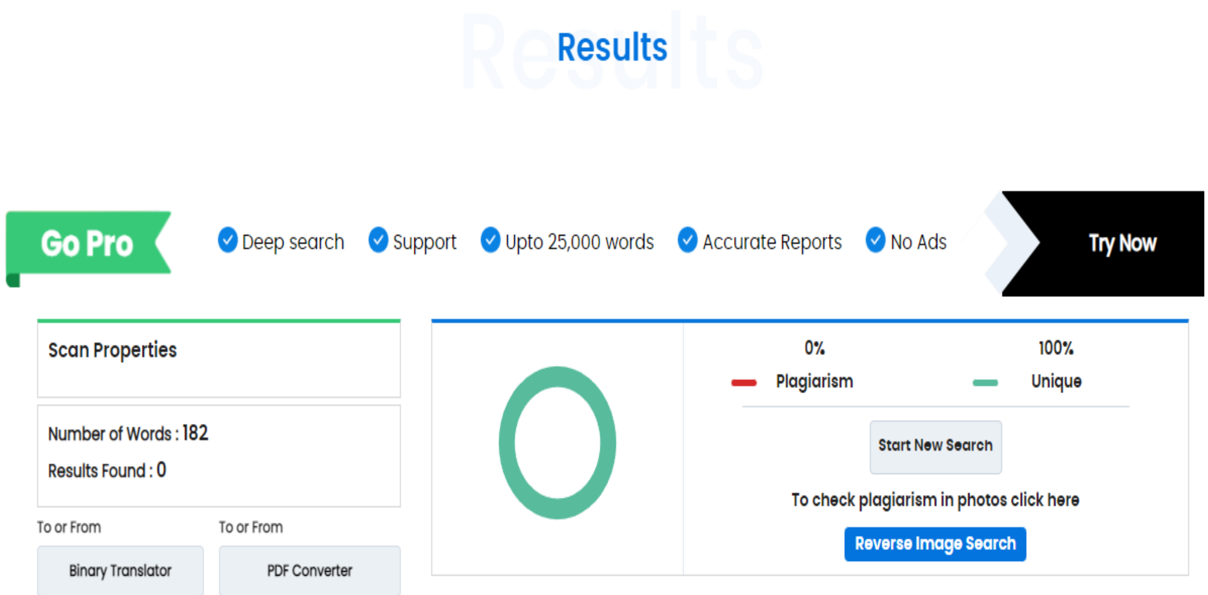
CHAPTER 2:



CHAPTER 3:



CHAPTER 4&5:


Results
Go Pro
Deep search
Support
Upto 25,000 words
Accurate Reports
No Ads
Try Now
Scan Properties
Number of Words : 182
Results Found : 0
To or From
Binary Translator
To or From
PDF Converter
0%
Plagiarism
100%
Unique
Start New Search
To check plagiarism in photos click here
Reverse Image Search