A MAIN PROJECT REPORT ON

CLASSIFICATION OF PNEUMONIA USING DEEP LEARNING

Submitted in partial fulfilment for the award of the degree of

BACHELOR OF TECHNOLOGY

In

Computer Science and Engineering

By

G.R.S Harshini (19A81A0514)

Under the Esteemed Supervision of

Mr. S. Kumar Reddy Mallidi M.Tech,(Ph.d)

Department of Computer Science and Engineering (Accredited by N.B.A.)

SRI VASAVI ENGINEERING COLLEGE(Autonomous)

(Affiliated to JNTUK, Kakinada)

Pedatadepalli, Tadepalligudem-534101, A.P 2022-23

SRI VASAVI ENGINEERING COLLEGE(Autonomous)

Department of Computer Science and Engineering (Accredited by N.B.A.)

Pedatadepalli, Tadepalligudem

Certificate

This is to certify that the Project Report entitled "**Classification of pneumonia using Deep Learning**" submitted by **G.R.S.Harshini (19A81A0514),** for the award of the degree of Bachelor of Technology in the Department of Computer Science and Engineering during the academic year 2022-2023.

Project Guide
Mr.S.Kumar Reddy Mallidi M.Tech., (PhD).
Assistant Professor

Head of the Department
Dr. D Jaya Kumari M.Tech., Ph.D.
Professor & HOD

External Examiner

DECLARATION

We hereby declare that the project report entitled "**CLASSIFICATION OF PNEUMONIA USING DEEP LEARNING**" submitted by us to Sri Vasavi Engineering College (Autonomous), Tadepalligudem, affiliated to JNTUK Kakinada in partial fulfilment of the requirement for the award ofthe degree of B.Tech in Computer Science and Engineering is a record of Bonafide project work carried out by us under the guidance of **Mr. S. Kumar Reddy Mallidi** **M.Tech,**(Ph.D). We further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of anyother degree in this institute or any other institute or University.

PROJECT ASSOCIATES

G. R. S. Harshini (19A81A0514)

ACKNOWLEDGEMENT

First and foremost, we sincerely salute to our esteemed institute **Sri Vasavi Engineering College** for giving this golden opportunity for fulfilling our warm dream to become engineer.

we consider it as our privilege to express our gratitude and respect to all those who guided inspired and helped us in completion of this project, we wish to express our sincere gratitude to our project guide, **Mr. S. Kumar Reddy Mallidi** M.Tech.,(PhD) and Assistant Professor and Dept. of Computer Science and Engineering, Sri Vasavi Engineering College, Pedatadepalli, for this timely cooperation and his valuable suggestions while carrying out this project.

We express our sincere thanks and heartful gratitude to **Dr. D. Jaya Kumari** Professor & HOD, Dept. of Computer Science and engineering, Sri Vasavi Engineering College, Pedatadepalli for permitting us to do our project.

We express our sincere thanks and heartful gratitude to **Dr. G. V. N. S. R. Ratnakar Rao**. Principal of Sri Vasavi Engineering College, Pedatadepalli for supporting us during this work assisted us in this project work. We would like to express our gratitude to our parents, friends those who helped to complete this course.

Project Associates

G.S.R.Harshini (19A81A0514)

ABSTRACT

Pneumonia is one of the major infectious diseases responsible for significant morbidity and mortality throughout the world. Having pneumonia can be frightening because this illness will impacts on functioning of lungs. It is a daunting task for the researchers to find suitable infection patterns on lung CT images for automated diagnosis of pneumonia. Classification of pneumonia from chest CT images based on reconstructed super-resolution images and VGG neural network. The problem with existing system is causing misclassification bias and mixtures of effects that threaten internal validity. A novel integrated semi-supervised shallow neural network framework comprising a SegNet for automatic segmentation of lung CT images followed by Fully Connected (FC) layers is proposed in our project. To overcome the drawback with existing system we are proposing a deep transfer learning model in which the input images are segmented and then classification will be done with high accuracy.

CHAPTER-1
INTRODUCTION

1.1Introduction

PNEUMONIA is an illness that disturbs the lung air sacs of an infected person. It is triggered by bacteria, fungi, or a virus that infects the air sacs of lungs that fill up with discharge fluids that leads to chills, fever, coughing with mucus, and breathing trouble among persons diagnosed with this disease. Children below five years of age and elderly patients with weak immune system are vulnerable to this type of diseases. Pneumonia has killed over a million children worldwide in 2018 and remains a life-threatening disease now a days if not detected or diagnose earlier.

1.2 Motivation

Segmentation of coronavirus disease 2019 (COVID-19) lesions is a difficult task due to high uncertainty in the shape, size and location of the lesions. CT scan image is an important means of diagnosing COVID-19, but it requires doctors to observe a large number of scan images repeatedly to determine the patient’s condition. Moreover, the low contrast of CT scan and the presence of tissues such as blood vessels in the background increase the difficulty of diagnosis. To solve this problem, we proposed an improved segmentation model called segNet model.

1.3 Scope

Radiograph of chest is penetrated through Scan where the soft tissues produce a dark color and hard tissues like bones produces a bright color. Patients diagnosed with pneumonia shows the chest cavity signs of fluids filling the air sacs of lungs as for the radiograph picture appears brighter. Several abnormalities may be seen on lung cavities as brighter color may represent such as cancer cells, blood vessels swelling, and abnormality of heart.

1.4 Objective

To validate the range and spot of an infected area of the lungs, chest scan is the utmost method. In this method, emergence of the disease can be imprecise and misinterpreted with another illness. Therefore, the undertaking is pleasing in the improvement of the processing in medical situations in isolated areas for pneumonia detection.

1.5 Project Outline

Chapter-1	Introduction
Chapter-2	Literature Survey
Chapter-3	System Study and Analysis
Chapter-4	System Design
Chapter-5	Technologies
Chapter-6	Implementation
Chapter-7	Result and Analysis
Chapter-8	Conclusion and Future Scope

CHAPTER-2
LITERATURE SURVEY

LITERATURE SURVEY

- Pneumonia is a common disease caused by different microbial species such as bacteria, virus, and fungi. The word “Pneumonia” comes from the Greek word “Pneumon” which translates to the lungs. Thus, the word pneumonia is associated with lung disease. In medical terms, pneumonia is a disease that causes inflammation of either one or both lung parenchyma. However, other causes of pneumonia include food aspiration and exposure to chemicals. Based on infection, pneumonia occurs as a result of inflammation caused by pathogens which lead the lung’s alveoli to fill up with fluid or pus and thereby leading to decrease of carbon dioxide (CO_2) and oxygen (O_2) exchange between blood and the lungs, making it hard for infected persons to breathe. The research work applied different Deep Learning algorithms for developing predictive model for pneumonia the main aim or prime goal of the analysis is to classify the algorithm that operates quicker, more reliably and more effectively in classification of pneumonia.

  However, it remains to investigate the performance of lung CT segmentation using the optimized version of PQIS-Net followed by classification with adaptive patch sizes. The authors are currently engaged in this direction.

- **He, X. et al. Sample- Effcient Deep Learning for COVID-19 Diagnosis Based on CT Scans. Preprint, Health Informatics (2020). https://doi.org/10.1101/2020.04.13.20063941**

  This study applied transfer learning on COVID-19 testing using CT images and discussed the impacts of various initialization parameters on the results, demonstrating that our model which was pre-trained on ImageNet21k has strong generalizability in terms of CT images. The proposed model provides an accuracy of 99.2% while detecting the COVID-19 cases. Compared to the neural architecture search model, our model shows the state of-the-art performance, across all metrics we have described. These ensure that COVID-19-negative patients are correctly diagnosed as negative in the vast majority of cases, reduce probability of diagnosing COVID19-negative cases as positive and reduce the burden on the health care system. Additionally, we examined the performance of the model with limited data and found that the model still performs satisfactorily. This shows that our model is still applicable with a limited data, which is characteristic of the real situation, where large and diverse datasets may not be readily available.

- **L. Fan, D. Li, H. Xue, L. Zhang, Z. Liu, B. Zhang, L. Zhang, W. Yang, B. Xie, X. Duan, X. Hu, K. Cheng, L. Peng, N. Yu, L. Song, H. Chen, X. Sui, N. Zheng, S. Liu, and Z. Jin, ‘‘Progress and prospect on imaging diagnosis of COVID-19,’’ Chin. J. Academic Radiol., vol. 3, no. 1, pp. 4–13, Mar. 2020, doi: 10.1007/s42058-020-00031-5.**

  In this work, a novel attempt has been made using an integrated semi-supervised shallow neural network encompassing the parallel self-supervised neural network model(PQIS-Net) for fully automatic segmentation of lung CT images followed by fully connected (FL) layers for patch-based classification with majority voting. The proposed integrated framework is semi-supervised in the sense that the PQIS-Net is a fully self-supervised network for segmentation followed by fully connected layers for random patch-based classification of COVID-19 disease.

The PQIS-Net model incorporates the frequency components of the weights and inputs in quantum formalism there by enabling faster convergence of the network states owing to reduction in computation. This intrinsic property of the PQIS-Net model yields precise and time efficient segmentation in real-time, which is evident from the results demonstrated in the experimental section. In spite being semi-supervised model, the suggested semi-supervised shallow neural network has attained better stability on the outcome as it is evident from the higher values of precision and AUC compared with the state-of-the-art methods specially in COVID-19 and Mycoplasma Pneumonia screening Moreover, our light-weight semi-supervised model can be employed in any application setting (eg .medical IoT devices) right away where, the deep learning models face serious obstacles.

Finally, we explored the relevant mechanism of COVID-19 testing using Grad-CAM visualization technique to make the proposed deep learning model more interpretable and explainable. The model performs performance validation through interpretability driven in a manner consistent with the radiologist's interpretation for the CP. The investigation of normal and NCP CT images helps to explore new visual indicators to assist clinical doctors in further manual screening. The experiments demonstrate that our models are effective in COVID-19 testing. In future, we will pay attention to the evaluation of the severity of COVID-19 and attempt to discover more valuable information from CT images to combat the pandemic. We will further conduct explanatory analyses on the models, which will shed light on the detection mechanism of COVID-19, to identify key characteristics in the CT images and to facilitate the screening by clinical doctors. Although the system has good performance on public datasets, the work is still at theoretical research stage, and the models have not been validated in actual clinical routine. Therefore, we will test our system in the clinical routine and communicate with physicians to understand how they use it and their opinions about the models. Thus, we can further improve the models in our future work.

CHAPTER-3

SYSTEM STUDY AND DESIGN

3.1 Problem Statement

Radiography, CT-scan or MRI is the common method to discover pneumonia. Medical personnel check the patient’s radiograph of their chest to determine if they are infected with pneumonia or not. In addition, the usual method for finding pneumonia is through medical history and laboratory results of the patient.

3.2 Existing System

Existing system developed a CNN model to detect common pneumonia disease from frontal and lateral chest Scan images. MIMIC-CXR dataset was used to perform large-scale automated recognition of these images. The dataset was split into training, testing and validation sets as 70%, 20% and 10%, respectively. Data augmentation and pixel normalization were used to improve overall performance. Their Dual Net CNN model achieved an average AUC of 0.72 and 0.688 for PA and AP, respectively.

3.3 Limitations

- Using X-Rays the classification of pneumonia is inaccurate.
- Medical diagnosis of patient is complicate
- It is time taking process.
- Less Efficiency even though using technology to determine the disease

3.4 Proposed System

This research explores the function in classification of pneumonia through deep learning using convolutional neural network (VGG19) for automatic segmentation of lung CT images followed by Fully Connected (FC) layers, is proposed in our project To overcome the drawback with existing system we are proposing a deep transfer learning model in which the input images are segmented through segNet and visualized through segNet algorithm and then classification will be done with high accuracy.

3.5 Advantages

- The use of segmentation adds colours to the affected area which classifies pneumonia more precisely.
- The proposed shape initialization method is robust to image noise and inhomogeneous intensity then learned shape dictionaries constrain the deformable model with shape priors such that it can effectively handle weak or missing nuclear boundaries.

3.6 Functional Requirement

In software engineering, a functional requirement defines a function of a software system or its component. A function is described as a set of inputs, the behavior, and outputs (see also software). Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioural requirements describing all the cases where the system uses the functional requirements are captured in use cases. Generally, functional requirements are expressed in the form “system shall do”. The plan for implementing functional requirements is detailed in the system design. In requirements engineering, functional requirements specify particular results of a system. Functional requirements drive the application architecture of a system. A requirements analyst generates use cases after gathering and validating a set of functional requirements. The hierarchy of functional requirements is: user/stakeholder request -> feature -> use case-> business rule. Functional requirements drive the application architecture of a system. A requirements analyst generates use cases after gathering and validating a set of functional requirements. Functional requirements may be technical details, data manipulation and other specific functionality of the project is to provide the information to the user.

The following are the Functional requirements of our system:

- The model selects the image from dataset.
- Detection can be done through Segmentation through segNet through which classifies the pneumonia.
- The result gives if the person is affected with viral pneumonia or bacterial pneumonia.

3.7 Non-Functional Requirements

In systems engineering and requirements engineering, a non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviours.

Usability:

Usability is a quality attribute that assesses how easy user interfaces are to use. The word "usability" also refers to methods for improving ease-of-use during the design process. Our project system is designed with completely automated process. Hence there is no or less user intervention.

Reliability:

Reliability is a quality or state of being reliable i.e; the extent to which an experiment, test or measuring procedure yields the same results on repeated traits. Our system is more reliable because of the quantities that are inherited from the chosen platform php. The code built by using php is more reliable.

Performance:

Performance is nothing but the execution of an action. As per our project the performance of our project can be described about how the system will give response. The system is developing in the high level language and using the advanced front-end and back-end technologies.it will give response to the end user on client system with in very less time.

3.8 System requirements

3.8.1 Hardware Requirements:

- System: Pentinum Dual Core.
- Hard Disk: 120 GB
- Monitor: 15”LED
- Input Devices: Keyboard, Mouse
- Ram: 1GB

3.8.2 Software Requirements:

- Operating System: Windows 7
- Coding Language: Python 3.x

CHAPTER-4

SYSTEM DESIGN

4.1 System Design

The proposed system is to classify the CT-scan images of chest into 5 types as viral Pneumonia, Bacterial pneumonia,c ovid-19,non covid viral and normal. In this proposed system we are mainly using machine learning algorithms. This research explores the deep learning function in detecting pneumonia through computer vision using convolutional neural network models. Every CNN models identified are assessed physically so that the framework segmentation and Grad-Cam visualization employed. The pneumonia infected diseases and normal chest scan image dataset are acquired from the Radiological Society of North America database. Our study enables to identify among the five models the best model to classify pneumonia. The CNN networks are the best model based on the observations of the researchers having an accuracy rate of 95% to 97%. The model has the lowest accuracy rate of 74%. All the models have performed well on classifying pneumonia and normal chest scans. With this work, a distinguished process of diagnosis and classification of pneumonia can helpful in providing medical services

4.2 System Architecture

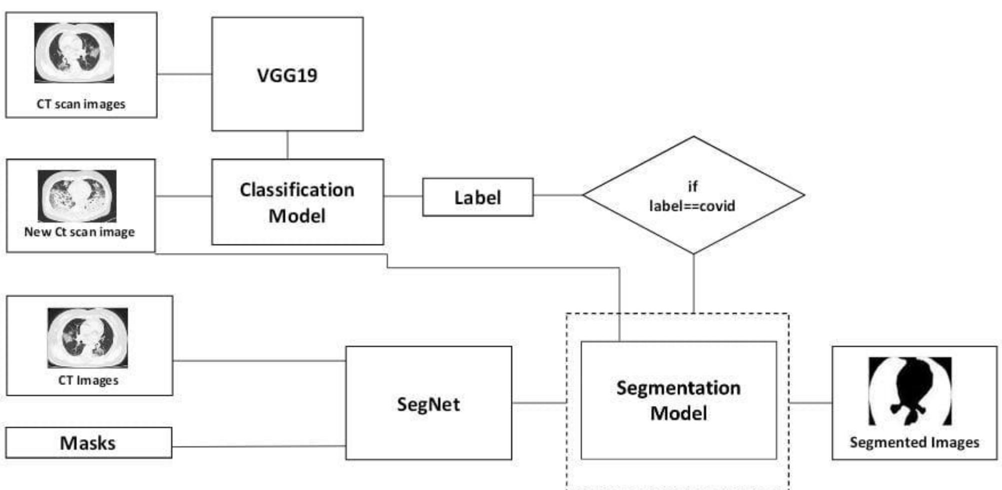
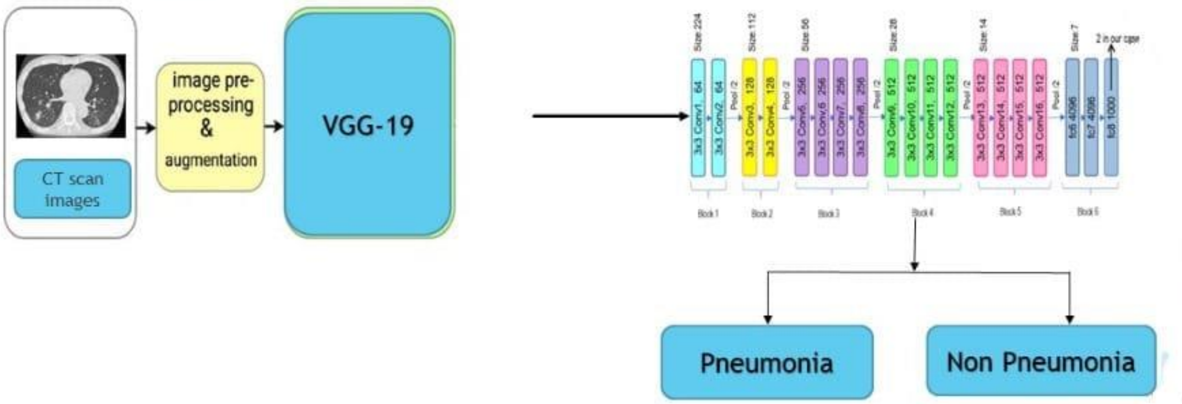


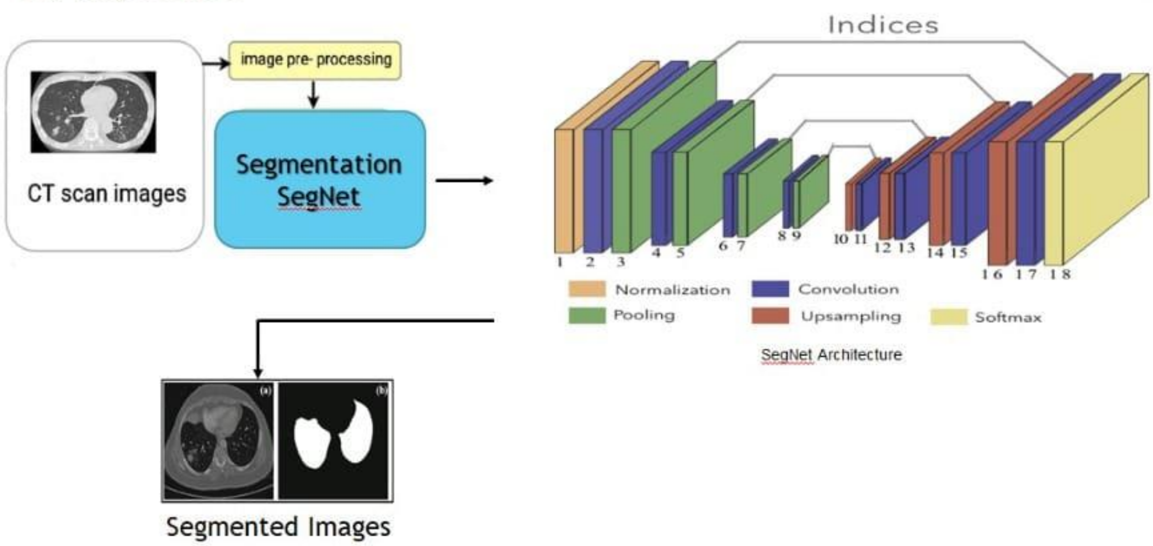
Fig 4.1 System Architecture

Our system is done in two tiers

Tier 1:Classification



Tier 2:Segmentation



Data Collection

Collection of data is from CT-scan images of chest. A data set is collected having chest CT-scan images on which these images are run in code undergoes segmentation.

CHAPTER-5

TECHNOLOGIES

5.1 Python

5.1.1 Python Introduction

Python is a general purpose, dynamic, high level and interpreted programming language. It supports an Object-Oriented programming approach to develop applications. It is simple and easy to learn and provides lots of high-level data structures. Python is easy to learn yet powerful and versatile scripting language which makes it attractive for Application Development. Python's syntax and dynamic typing with its interpreted nature, makes it an ideal language for scripting and rapid application development. Python supports multiple programming pattern, including object oriented, imperative and functional or procedural programming styles. Python is not intended to work on special areas such as web programming.

5.1.2 Python History

Python laid its foundation in the late 1980s.

1) The implementation of Python was started in December 1989 by Guido Van Rossum at CWI in the Netherlands.
2) In February 1991, van Rossum published the code (labelled version 0.9.0).
3) In 1994, Python 1.0 was released with new features like: lambda, map, filter, and reduce.
4) Python 2.0 added new features like: list comprehensions, garbage collection system.
5) On December 3, 2008, Python 3.0 (also called "Py3K") was released. It was designed to rectify fundamental flaws of the lang

5.1.3 Python Features

Python provides lots of features that are listed below.

1) Easy to Learn and Use Python is easy to learn and use. It is a developer-friendly and high-level programming language.
2) Expressive Language Python language is more expressive means that it is more understandable and readable.
3) Interpreted Language Python is an interpreted language i.e., interpreter executes the code line by line at a time. This makes debugging easy and thus suitable for beginners.

4) Cross-platform Language Python can run equally on different platforms such as Windows, Linux, Unix and Macintosh etc. So, we can say that Python is a portable language.
5) Free and Open-Source Python language is freely available at official web address. The source-code is also available. Therefore, it is open source.
6) Object-Oriented Language Python supports object-oriented language and concepts of classes and objects come into existence.
7) Extensible it implies that other languages such as C/C++ can be used to compile the code and thus it can be used further in our python code.
8) Large Standard Library Python has a large and broad library and provides a rich set of modules and functions for rapid application development.
9) GUI Programming Support Graphical user interfaces can be developed using Python.
10) Integrated IT can be easily integrated with languages like C, C++, JAVA etc.

CHAPTER-6

IMPLEMENTATION

6. Implementation Steps

- Here we are using python language. .
- To compile the code we are using Google co-lab.
- To create the columns for output we are using tensor flow ,numpy ,matplotlib,keras libraries.
- Here we are implementing

Data collection

Segmentation

Display of output

6.1Algorithms used:

- Convolutional Neural Network (CNN):-
- A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.
- CNNs are used for image classification and recognition because of its high accuracy. The CNN follows a hierarchical model which works on building a network, like a funnel, and finally gives out a fully-connected layer where all the neurons are connected to each other and the output is processed.

In this work a novel attempt has been made using the segmentation (segNet model) encompassing the convolutional neural network model(VGG19) for classification of pneumonia disease.

Classification of pneumonia

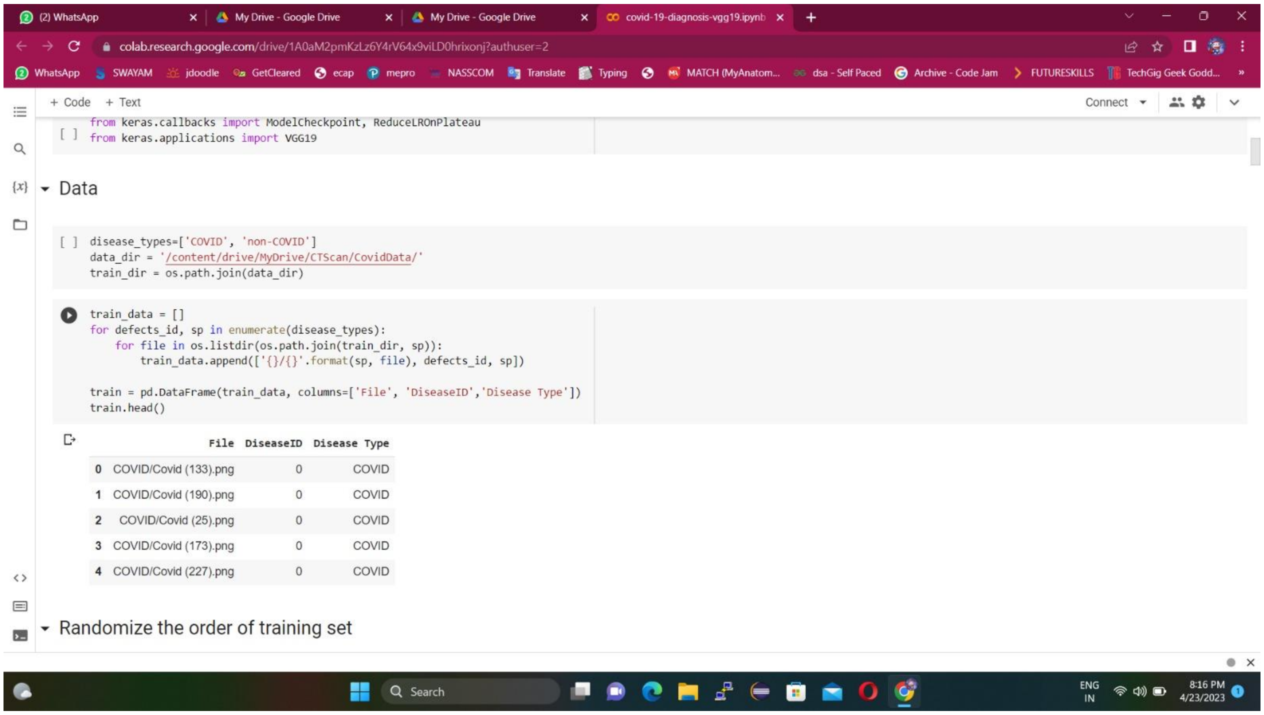


Fig: 6.1 training the data

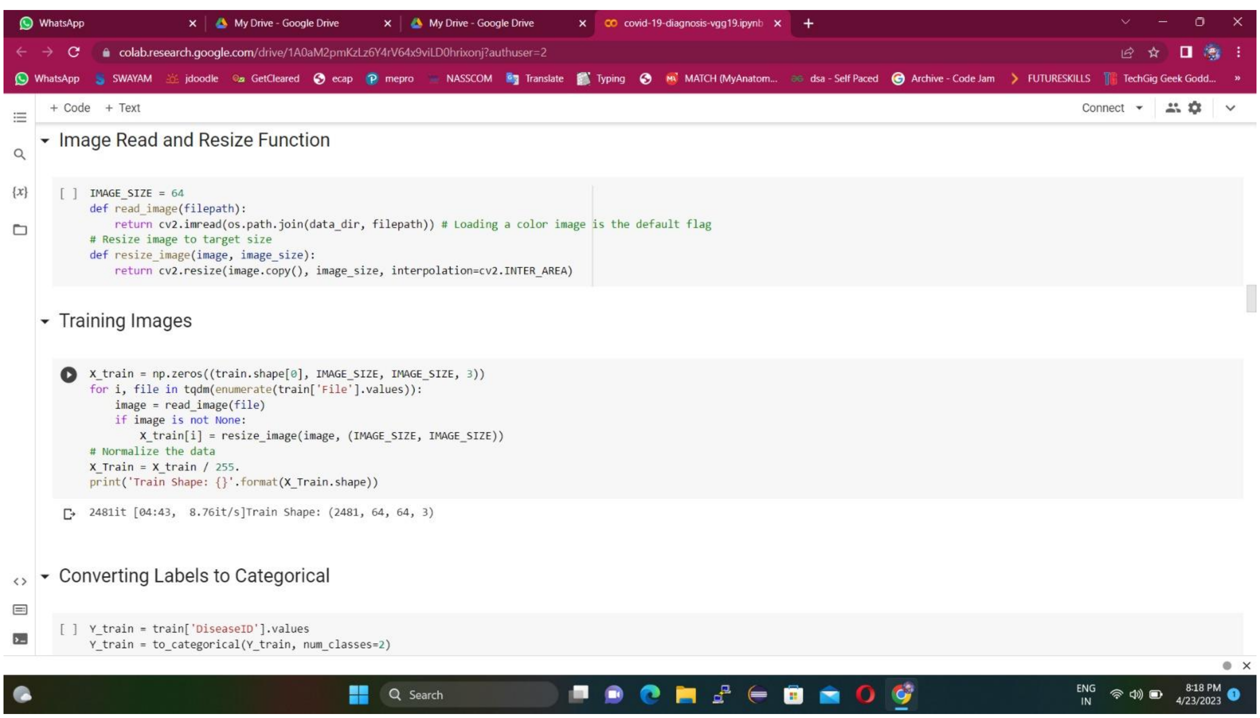


Fig: 6.2 reading images and resizing

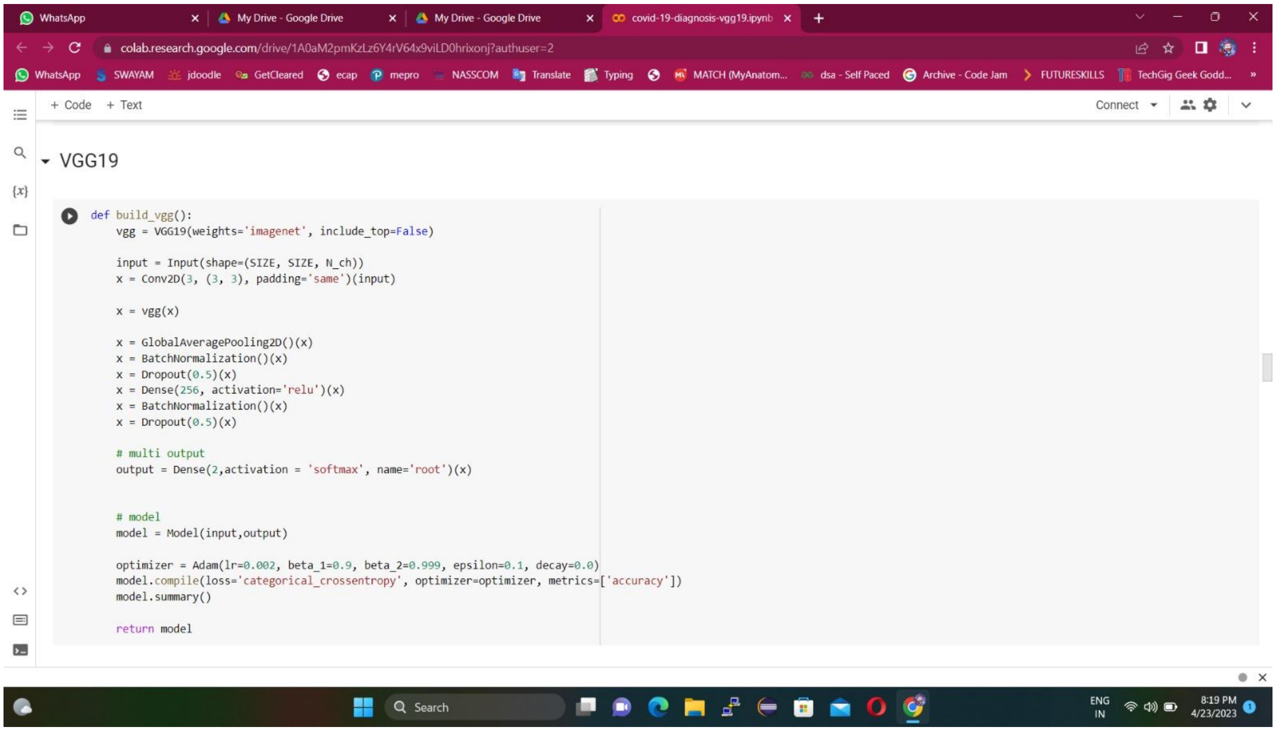


Fig 6.3 : Implementation of VGG 19

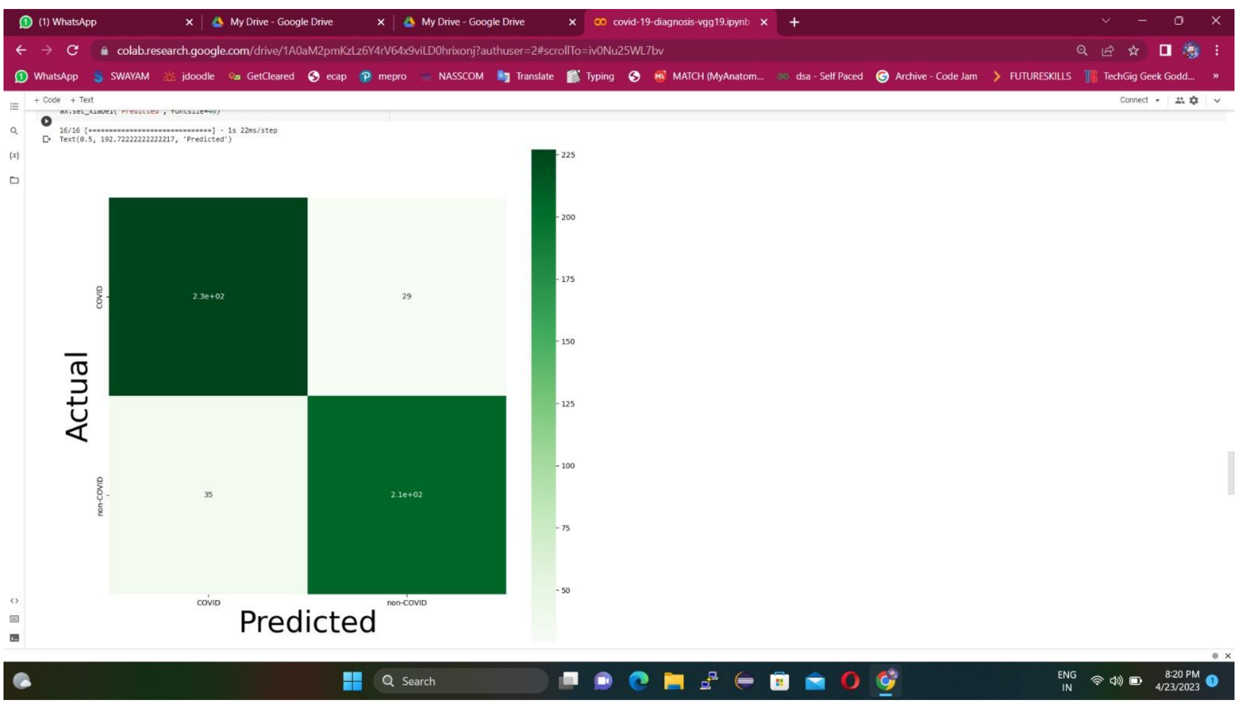


Fig 6.4 : confusion matrix

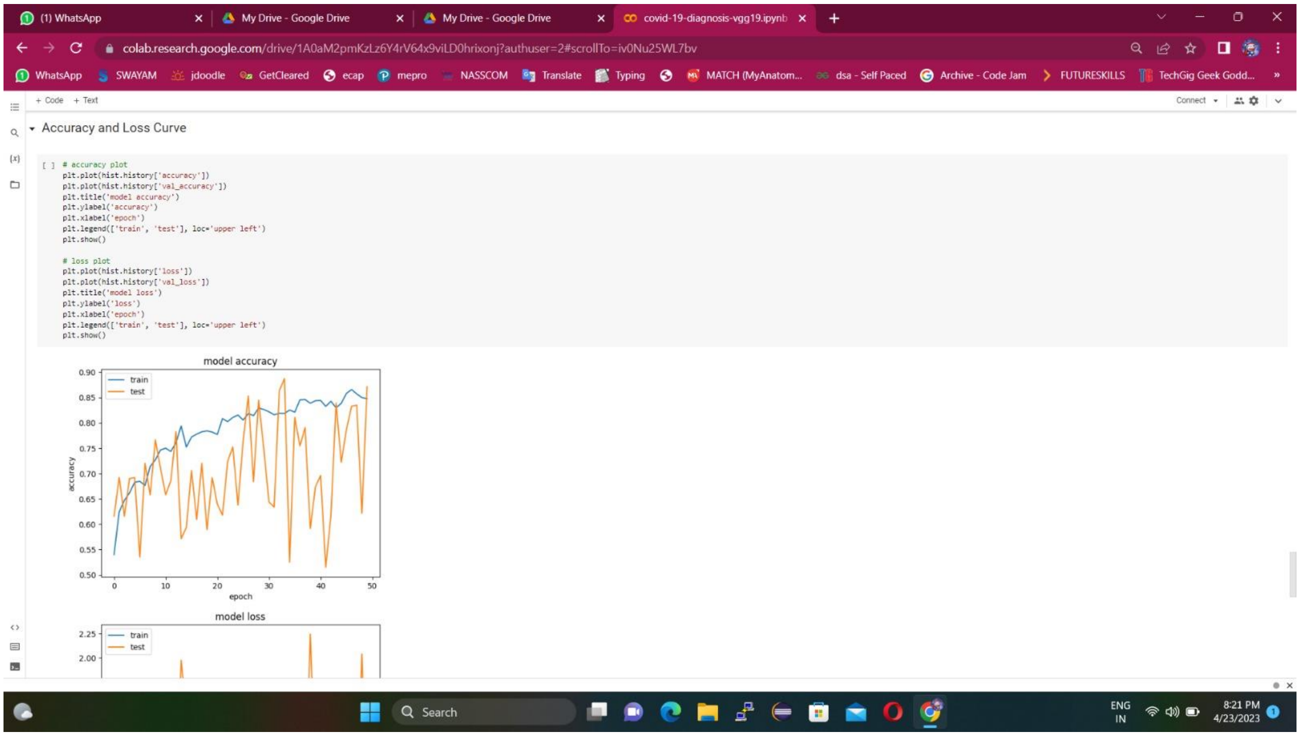


Fig 6.5: Evaluating the Accuracy and Loss curve

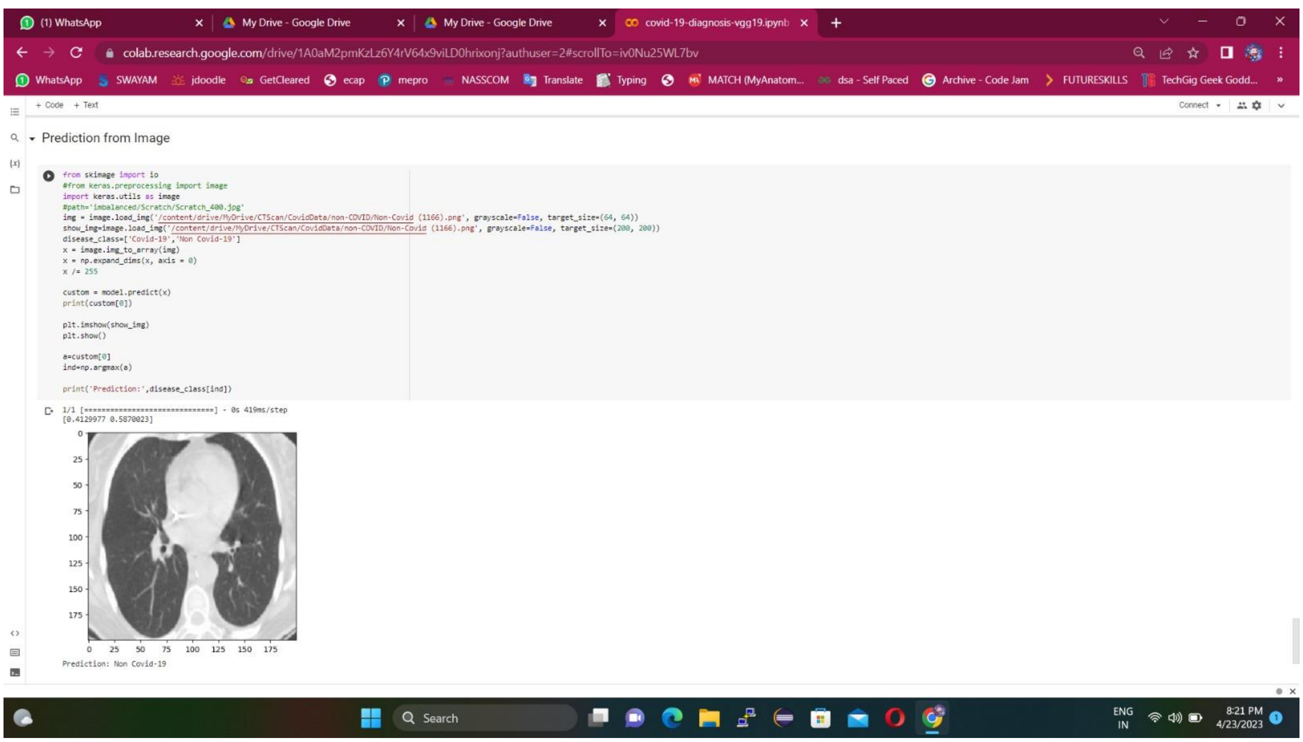
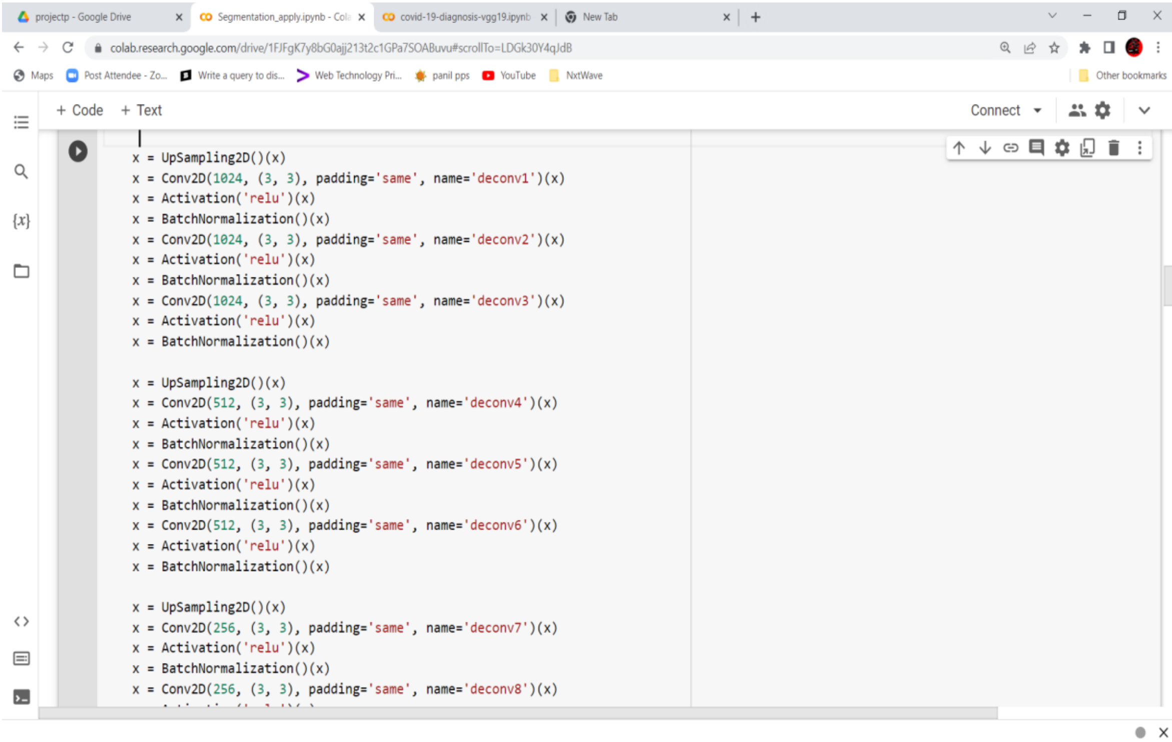


Fig 6.6 : prediction of result from image segmentation

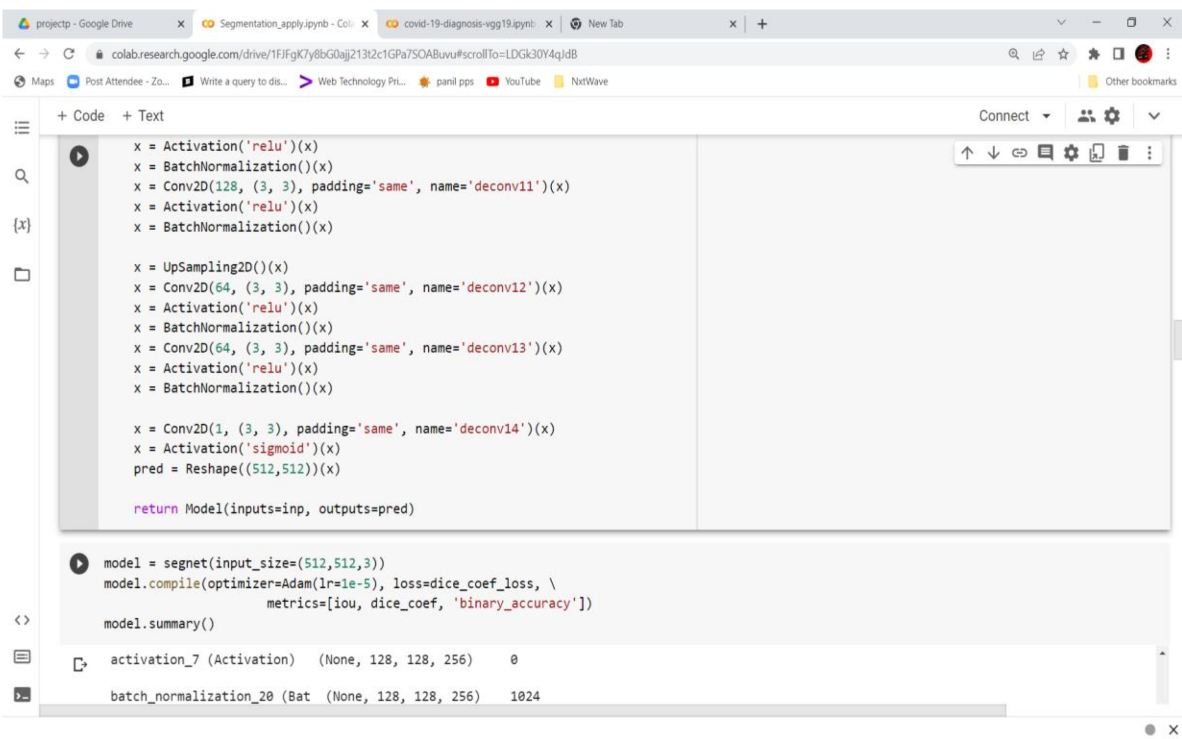


```
x = UpSampling2D()(x)
x = Conv2D(1024, (3, 3), padding='same', name='deconv1')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(1024, (3, 3), padding='same', name='deconv2')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(1024, (3, 3), padding='same', name='deconv3')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)

x = UpSampling2D()(x)
x = Conv2D(512, (3, 3), padding='same', name='deconv4')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(512, (3, 3), padding='same', name='deconv5')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(512, (3, 3), padding='same', name='deconv6')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)

x = UpSampling2D()(x)
x = Conv2D(256, (3, 3), padding='same', name='deconv7')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(256, (3, 3), padding='same', name='deconv8')(x)
```

Fig:6.7 Inputs for segnet model



```
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(128, (3, 3), padding='same', name='deconv11')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)

x = UpSampling2D()(x)
x = Conv2D(64, (3, 3), padding='same', name='deconv12')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)
x = Conv2D(64, (3, 3), padding='same', name='deconv13')(x)
x = Activation('relu')(x)
x = BatchNormalization()(x)

x = Conv2D(1, (3, 3), padding='same', name='deconv14')(x)
x = Activation('sigmoid')(x)
pred = Reshape((512,512))(x)

return Model(inputs=inp, outputs=pred)
```

```
model = segnet(input_size=(512,512,3))
model.compile(optimizer=Adam(lr=1e-5), loss=dice_coef_loss, \
                    metrics=[iou, dice_coef, 'binary_accuracy'])
model.summary()
```

```
activation_7 (Activation)    (None, 128, 128, 256)    0

batch_normalization_20 (Bat  (None, 128, 128, 256)    1024
```

Fig 6.8: VGG 19

CHAPTER-7 RESULT & ANALYSIS

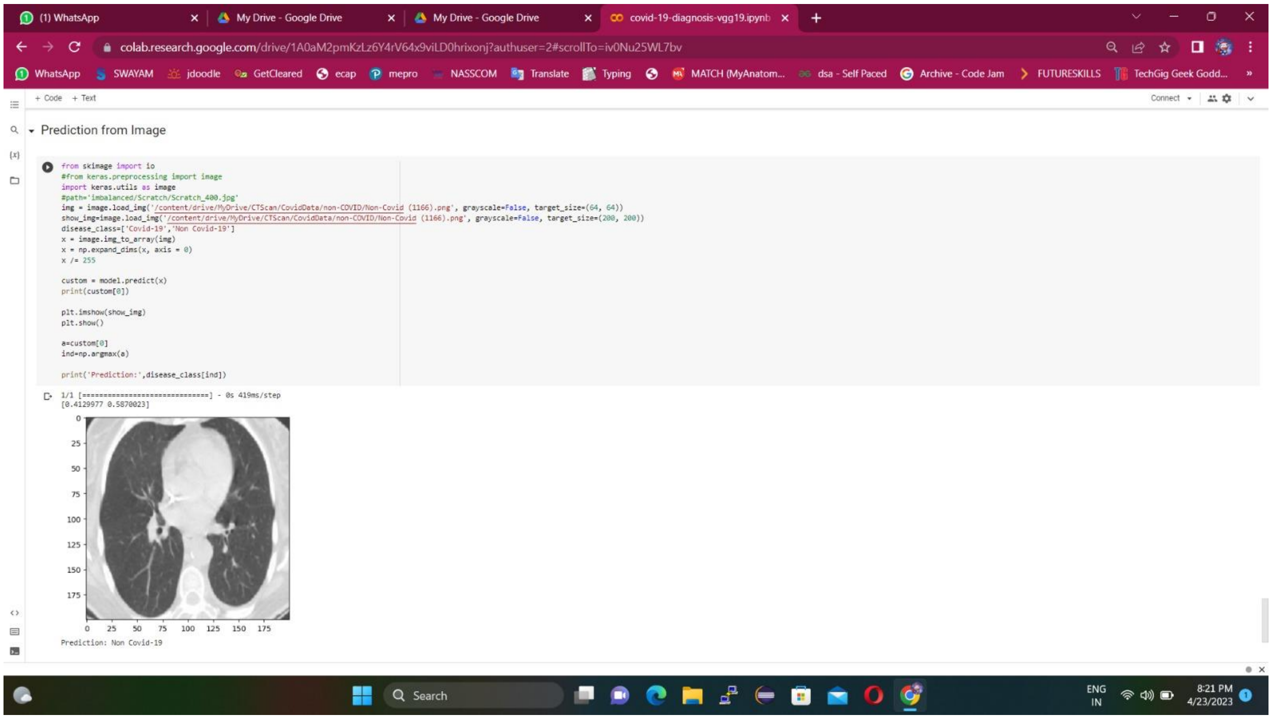


Fig 7.0: Prediction of result from image

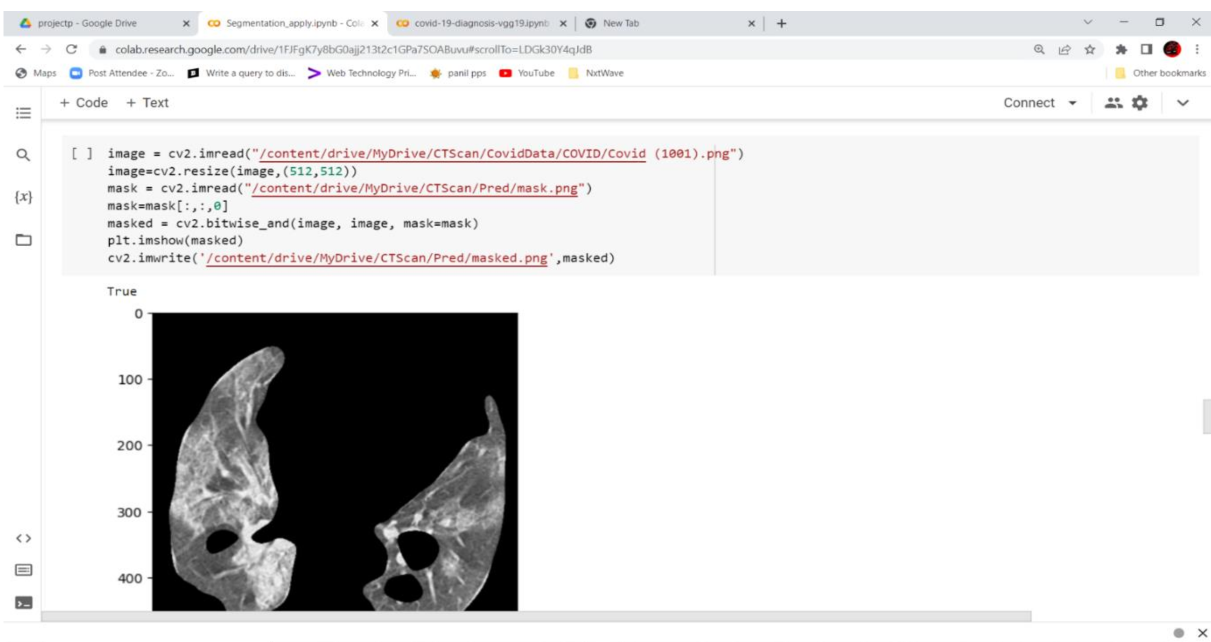


Fig7.1: displaying input for segnet model

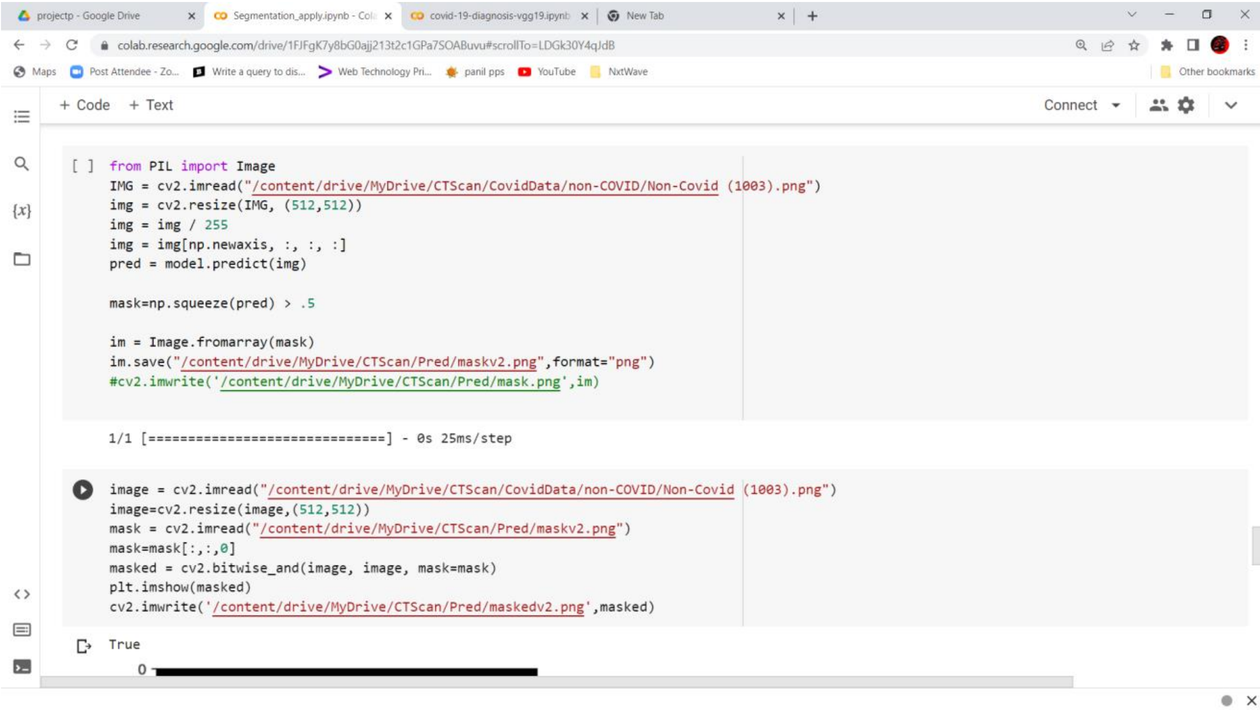


Fig:7.2: Prediction of effected area from the image

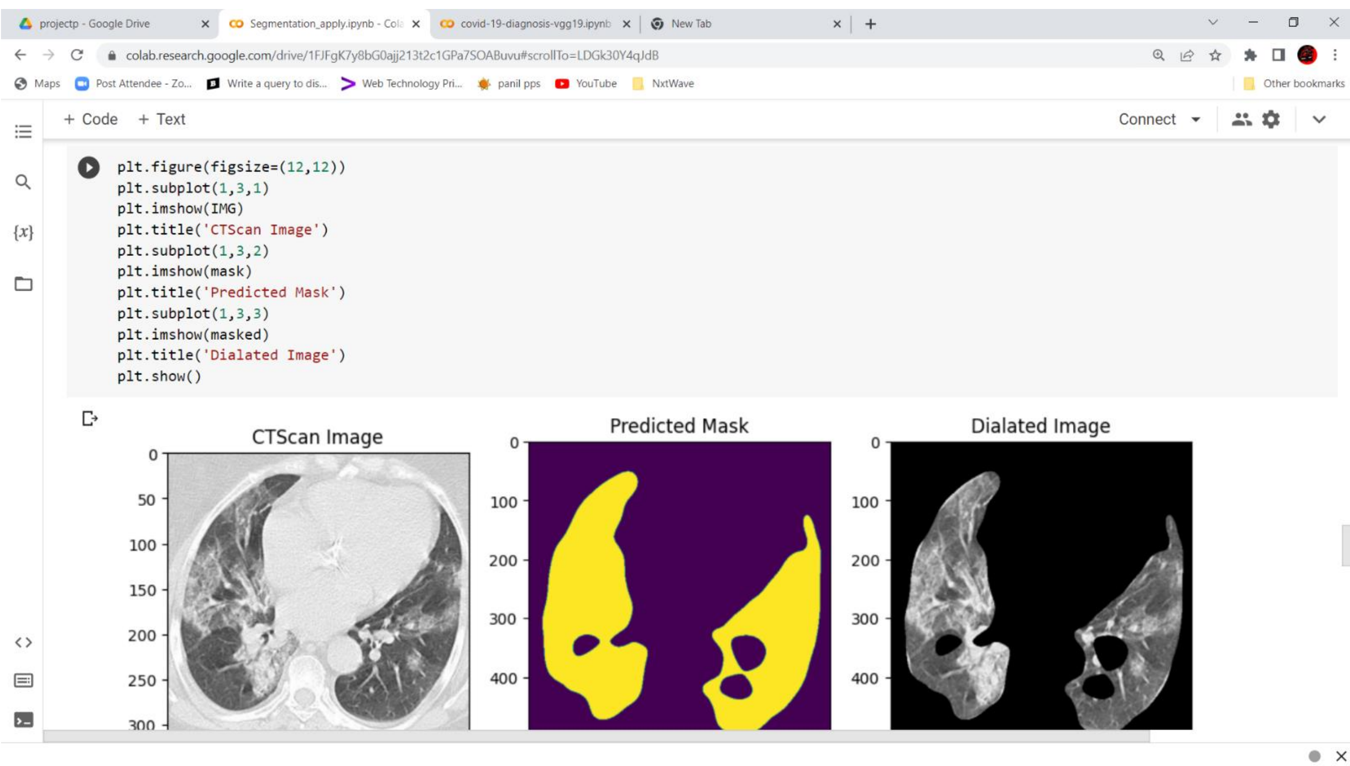


Fig:7.3: Applying the mask for predicted area

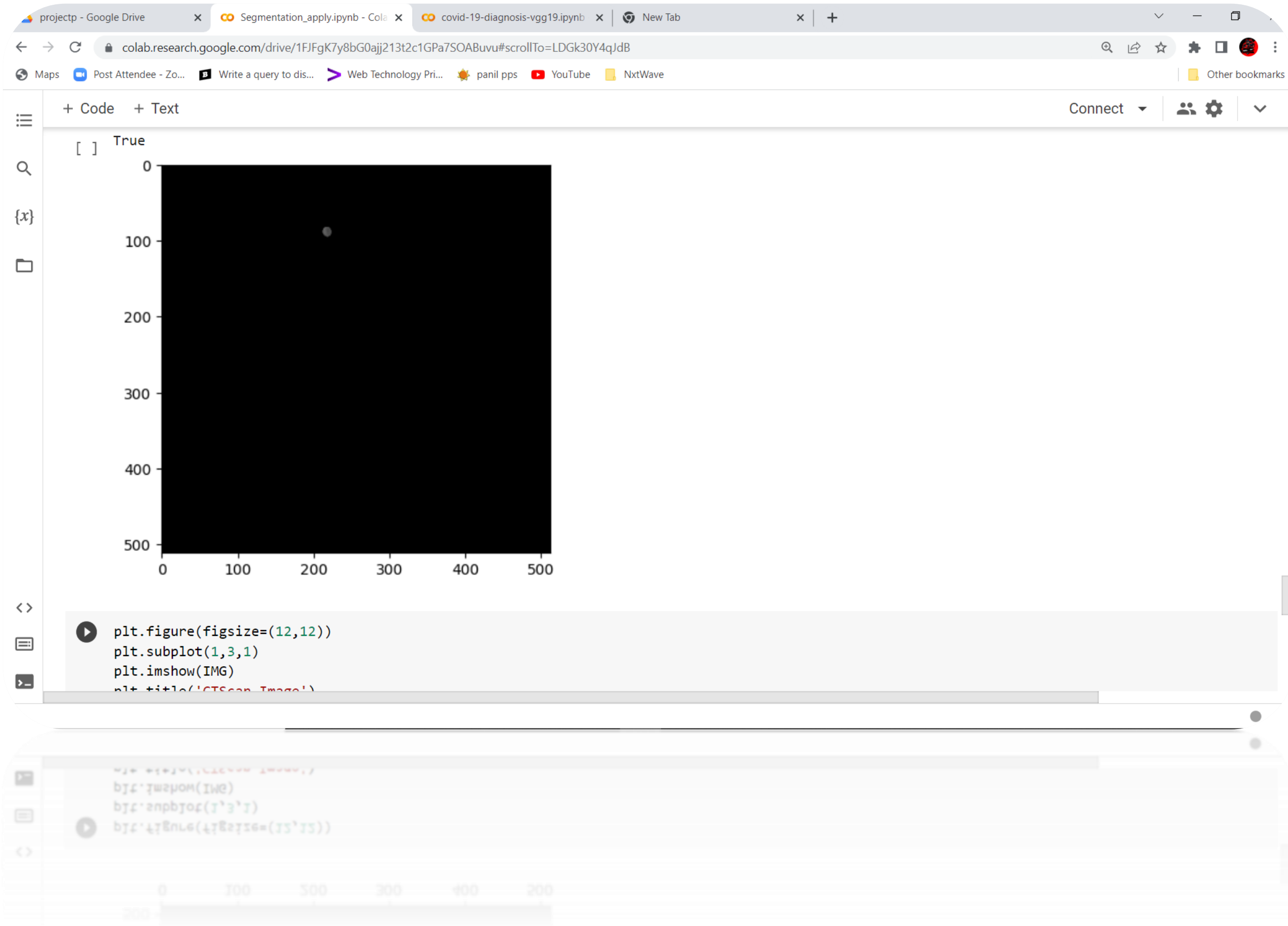


Fig:7.4: Applying the mask for predicted area

CHAPTER-8

CONCLUSION & FUTURE WORK

Conclusion and future work:

- The final conclusion of our project is we developed a model to classify the pneumonia using Deep learning. In this model it takes CT Scan images as input which are pneumonia or non pneumonia and classify whether the CT Scan image belongs to pneumonia patient or non pneumonia patient. To develop our model we have used CNN algorithms VGG-19 in the techniques of segmentation and segNet. Finally we get output about the given image is either covid or non pneumonia.

Future Work:

- The future scope of pneumonia classification using machine learning algorithms is promising, as advancements in technology and data analysis continue to improve the accuracy, speed, and scalability of pneumonia classification systems.
- The pneumonia further can be classified as either viral or bacterial pneumonia.

REFERENCES:

1. D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction", *2019 IEEE International Conf. on Electrical Computer and Communication Technologies (ICECCT)*, vol.395,pp. 1-7, 2019.
2. Stephen Okeke et al., "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare", *Journal of healthcare engineering 2019*, *Cell*, vol. 172.5, pp. 1122-1131, 2018.
3. He, X. et al. Sample- Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans. Preprint, Health Informatics (2020).
4. L. Fan, D. Li, H. Xue, L. Zhang, Z. Liu, B. Zhang, L. Zhang, W. Yang, B. Xie, X. Duan, X. Hu, K. Cheng, L. Peng, N. Yu, L. Song, H. Chen, X. Sui, N. Zheng, S. Liu, and Z. Jin, ‘‘Progress and prospect on imaging diagnosis of COVID-19,’’ Chin. J. Academic Radiol., vol. 3, no. 1, pp. 4–13, Mar. 2020, doi: 10.1007/s42058-020-00031-5.
5. Nicholas E Ross, Charles J Pritchard, David M Rubin and Adriano G Duse, "Automated image processing method for the diagnosis and classification of pneumonia ", Medical and Biological Engineering and Computing, vol. 44, no. 5, 2006.