**CLINICAL DESCISION SUPPORT SYSTEM FOR COVID-19**

A PROJECT REPORT

Submitted by

**Garlapati Pavan Kumar**

Reg. No.**16MIS1097**

in partial fulfillment for the award of the degree of

Master of Technology

in

Software Engineering (5 Year Integrated Programme)





**School of Computer Science and Engineering**

Vellore Institute of Technology

Vandalur - Kelambakkam Road, Chennai - 600 127

June - 2021



School of Computer Science and Engineering

DECLARATION

I hereby declare that the project entitled Clinical Decision Support System for Covid-19 submitted by me to the School of Computer Science and Engineering, VIT Chennai, 600 127 in partial fulfillment of the requirements of the award of the degree of Master of Technology in Software Engineering (5 year Integrated Programme) is a bona-fide record of the work carried out by me under the supervision of Prof.Shola Usha Rani. I 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 any other degree or diploma of this institute or of any other institute or University.

Place: Chennai

Date:

Signature of Candidate



School of Computer Science and Engineering

CERTIFICATE

This is to certify that the report entitled Clinical Decision Support System is prepared and sub-mitted by Garlapati Pavan Kumar (Reg. No.16MIS1097) to VIT Chennai, in partial fulfullment of the requirement for the award of the degree of Master of Technology in Software Engineering (5 year Integrated Programme) is a bona-fide record carried out under my guidance. The project fulfills the requirements as per the regulations of VIT and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

Guide/Supervisor HoD

Name:Prof.Shola Usha Rani Name: Dr. Asnath Victy Phamila Y

Date: Date:

Examiner Examiner

Name: Name:

Date: Date:

(Seal of SCSE)

iv

Acknowledgement

I obliged to give my appreciation to a number of people without whom I could not have completed this thesis successfully.

I would like to place on record my deep sense of gratitude and thanks to my internal guide Prof.Shola Usha Rani , School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Chennai, whose esteemed support and immense guidance encouraged me to complete the project successfully.

I would like to thank our HoD Dr. Asnath Victy Phamila Y, School of Computer Science and Engineering (SCOPE) and Project Co-Ordinator Dr. C. Sweetlin Hemalatha, Vellore Institute of Technology, Chennai, for their valuable support and encouragement to take up and complete this thesis.

Special mention to our dean Dr. Jagadeesh Kannan R , Associate Dean Dr. S. Geetha, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Chennai, for motivating us in every aspect of software engineering.

I thank our management of Vellore Institute of Technology, Chennai, for permitting me to use the library and laboratory resources. I also thank all the faculty members for giving me the courage and the strength that I needed to complete my goal. This acknowledgment would be incomplete without expressing the whole hearted thanks to my family and friends who motivated me during the course of my work.

Garlapati Pavan Kumar

Reg. No. 16MIS1097

v

Abstract:

Currently, Coronavirus disease (COVID-19),one of the most infectious diseases in the 21st century, is diagnosed using RT-PCR testing, CT scans and/or Chest X-Ray images. CT (Computed Tomography) scanners and RT-PCR testing are not available in most medical centers and hence in many cases Chest X-ray images become the most time/cost effective tool for assisting clinicians in making decisions.

Deep learning neural networks have a great potential for building COVID-19 triage systems and detecting COVID-19 patients.My Project aims to build a Deep Learning Method that can easily/rapidly identify and visualize the opacity which will help the medical personal to swiftly give diagnosis.

vi

Contents

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Declaration | | |  | i |
|  | Certificate | | |  | ii |
|  | Acknowledgement | | | | iii |
|  | Abstract | | |  | iv |
| 1 | Introduction | | |  | 1 |
|  | 1.1 | Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | | 1 |
|  | 1.2 | Problem Statement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | |  |
|  | 1.3 | Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | |  |
|  | 1.4 | Challenges . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | | |  |
| 2 Planning & Requirements Specification | | | | |  |
|  | 2.1 | System Planning . . . . . . . . . . . . . . . . . . . . . . . . . . | | |  |
|  | 2.2 | System Requirements . . . . . . . . . . . . . . . . . . . . . . . . | | |  |
|  |  | 2.2.1 | | Hardware Requirements . . . . . . . . . . . . . . . . . . |  |
|  |  | 2.2.2 | | Software Requirements . . . . . . . . . . . . . . . . . . . |  |
| 3 | System Design | | | |  |
|  |  | | | |  |
| 4 | Implementation of System | | | |  |
|  |  |  | | |  |
| 5 | Results & Discussion | | | |  |
|  |  |  | | |  |
| 6 | Conclusion and Future Work | | | |  |
|  |  |  |  | |  |
|  | References | | | |  |
|  | Appendix <Sample code, snapshot etc.> | | | |  |

vii

List of Tables

|  |  |  |
| --- | --- | --- |
| S.No | Table Name | Page No. |
| 1 | List of Figures | 8 |
| 2 | Introduction | 9 |
| 3 | Results & Discussion | 20 |

viii

List of Figures

|  |  |  |
| --- | --- | --- |
| S.No | Description | Page No |
| 1.1 | FOUR PHASE DESIGN | 10 |
| 1.2 | Design Diagram | 13 |
| 1.3 | Data Flow Diagram | 14 |
| 1.4 | Doctor use Case | 15 |
| 1.5 | Patient Use Case | 16 |
| 1.6 | Alex-Net Architecture Diagram | 17 |
| 1.7 | Alex-Net Architecture | 18 |
| 1.8 | Data-set Split Diagram | 19 |
| 1.9 | Data-set Split code | 19 |
| 1.10 | Random Images after RGB | 20 |
| 1.11 | Model Accuracy Graph | 21 |
| 1.12 | Cross Entropy | 21 |
| 1.13 | Confusion Matrix | 22 |
| 1.14 | Precision, Recall, F-1 Scores of Each Class | 22 |
| 1.15 | Model Prediction | 23 |
| 1.16 | Grad Cam for Normal | 23 |
| 1.17 | Grad Cam for COVID | 24 |
| 1.18 | Grad Cam for Lung Opacity | 24 |
| 1.19 | Grad Cam for Viral Pneumonia | 25 |

ix

1.Introduction

* 1. BACKGROUND

The Coronavirus has Devestated the World today and Doctors are scarce as does the Diagnosis.All the medical professional around the world are in a Desperate Need for Intellegent Machine that will be Efficient, Accurate, Fast Decison Making & that could Visually able to identify the Opacity.

* 1. PROBLEM STATEMENT

Even with the latest tech-advancements the “Lack of Information and Clinical Decision Making” are becoming difficult to health caretakers. Especially during these days the efficient decision making is necessary as the patient situation gets very adverse in narrow time.

* 1. MOTIVATION

The Motivation of the system is to serve as a decision-making platform for the health care professionals as its rapidly spreading and flow of inpatients is high.

* 1. CHALLENGES

1. To achieve the optimum accuracy is one of the great challenges.

2. As the data related to X-rays are very confidential and very hard to acquire.

3. “High Precession & Fast Assitance” for medical professionals.

4. Aggregation of the high-grade DataSet.

x

1. Planning & Requirements Specification

2.1 System Planning

Clinical Decision making aims to develop a Dense CNN to predict whether the input

image is having COVID-19 so for this high-grade data sets, desired set of knowledge

need to be obtained and it will be achieved in sequential steps.

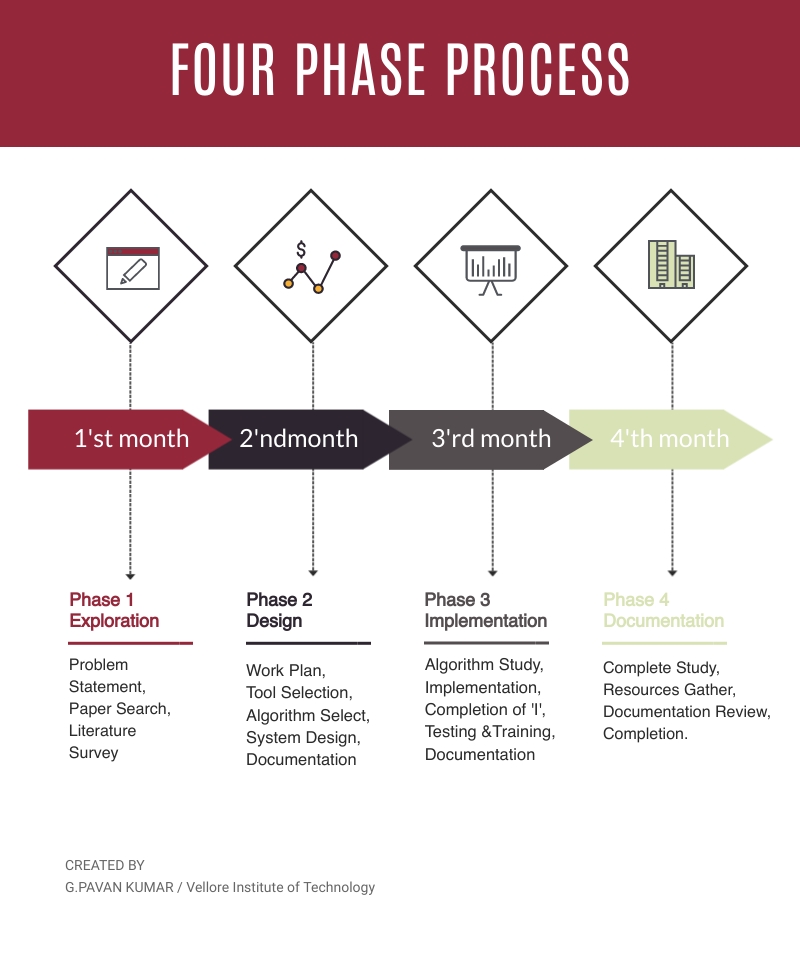


Figure 1.1 FOUR PHASE DESIGN

Phase-1:Exploration

Aims to identify problem and perform a deep search for literature survey and research paper.

Phase-2: Design

In this phase the objectives are identified based on that a work plan is designed and based on the research paper we have to identify a efficient algorithm and document the results above

xi

Phase-3&4: Implementation

A regressive and professional algorithmic study need to be performed so that to implement a better algorithm later stages includes training and testing and completion the complete documentation.

xii

2.2 System Requirements

2.2.1 Hardware Requirements

1) Processor: Intel Pentium IV 2.0 GHz and above.

2) RAM: 512 MB and Above.

3) Hard Disk: 80GB and Above.

4) Monitor: CRT or LCD monitor.

5) Keyboard: Normal or Multimedia.

6) Mouse: Any

2.2.2 Software Requirements

1. Language: Python3.5 or Above .

2. OS: Windows 7 or above.

3. Tool: Google Colaboratory OR Kaggle.

4.Packages:matplotlib=3.3.4, numpy=1.20.2, Pillow=8.0.1, Open-cv=4.5.1.48, Seaborn =0.11.1,torch=1.8.0, torchvision=0.9.0, tqdm=4.60.0,Tensorflow=1.2(above), Sci-kit, itisdangerous.

5.Processing Unit: TPU-18 and GPU.

xiii

1. System Design

DESIGN DIAGRAM:

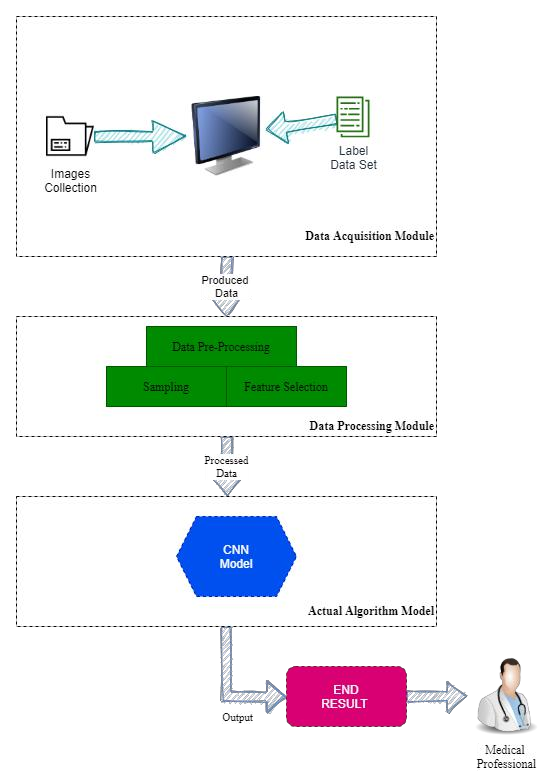


Figure1.2 Design Diagram

1.Data Acquisition Module:

In this module the Chest CT/X-ray images are collected from various open sources with their label data.

2.Data Preprocessing Module:

In this module the sampling and feature selection are done using the label data of the images.

3.Actual Algorithm Module:

Appropriate algorithm is programmed to identify the opacities in the chest

CT/X-ray image.

xiv

DATA FLOW DIAGRAM

The below Data Flow diagram explains how the data is processed in each stage and what happens in between



Figure 1.3 Data Flow Diagram

xv

USE CASE DIAGRAMS:

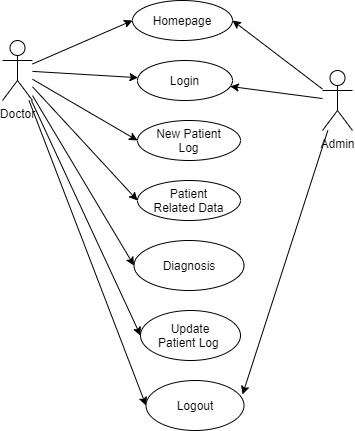


Figure 1.4 Doctor Use Case:

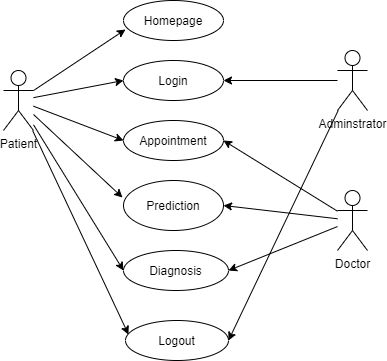


Figure 1.5 Patient Use Case

xvi

4.Implementation of System

EXISTING SYSTEM**:**

RT-PCR(Reverse transcription polymerase chain reaction):

* This is a lab technique in which multiple copies of RNA are made and then transcript-ed into DNA.
* Currently Rt-Pcr System is in use to confirm whether a patient is positive or not.

Disadvantages:

* The method is very sensitive, even a tiny low quantity of polymer contamination will cause false results.
* Trained Experts and high end equipment needed for such testing
* The distribution of virus across the tract varies between patients, therefore albeit an individual is infected, the virus might solely be detectable in sputum or nasal swab however not essentially at each locations at an equivalent time.
* This is time taking and not affordable by everyone and due to many tests need to be conducted accurate identification will be low.

PROPOSED SYSTEM**:**

The system aims to investigate the opacity features for detecting the Covid-19 using the Alex-Net and in later stages they are tested and trained, visualization are implemented for efficient results.After that it is saved into .h5 file which will be used to deploy a front end application .In the next stages a the graphs are developed and target image is given as input to get prediction with Grad-Cam Visualization.

Advantages**:**

* The Proposed System is developed in such way to overcome the above disadvantages can be effectively and maintaining the higher accuracy
* The proposed system is built using Alex-Net CNN which is generally opted for models using medical imaging.
* The Data Sets used for the proposed system are gathered from open sources mainly to encourage Research and further deep study.
* Instead of opting for expensive and ineffective practices the model provides to detect covid-19 speedily and effectively.

xvii

* Proposed system doesn’t need for any trained expertise or high end equipment.

Alex-Net CNN:

It’s name came from one of the proficient authors- AlexKrizhevsky and it is one of the prime deep convolutional neural networks developed in 2012.AlexNet has a similar structure to that of LeNet,but uses more convolutional layers and a larger parameter space to fit the large-scale ImageNet dataset.Today AlexNet has been surpassed by much more effective architectures but it is a key step from shallow to deep networks that are used nowadays.

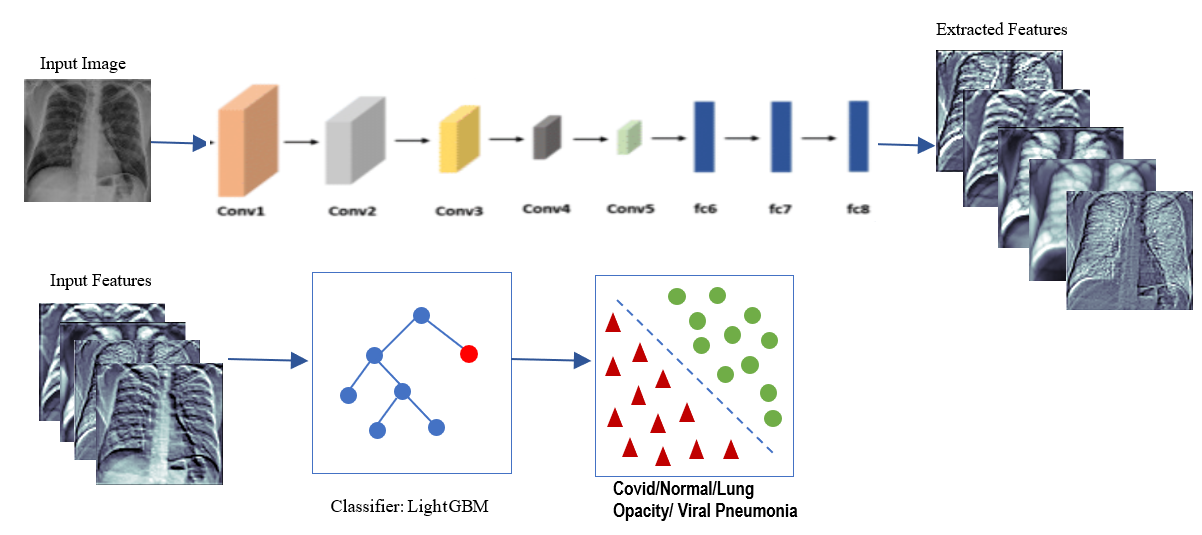


Figure 1.6: Alex-Net Architecture Diagram

Architecture:

Alex-Net consists of total 8 layers from which 5 are convolutional layers with 1,2,5 as max-pooling layers with excellent and proper feature extraction and later that three fully connected layers each with Relu and dropout.

xviii

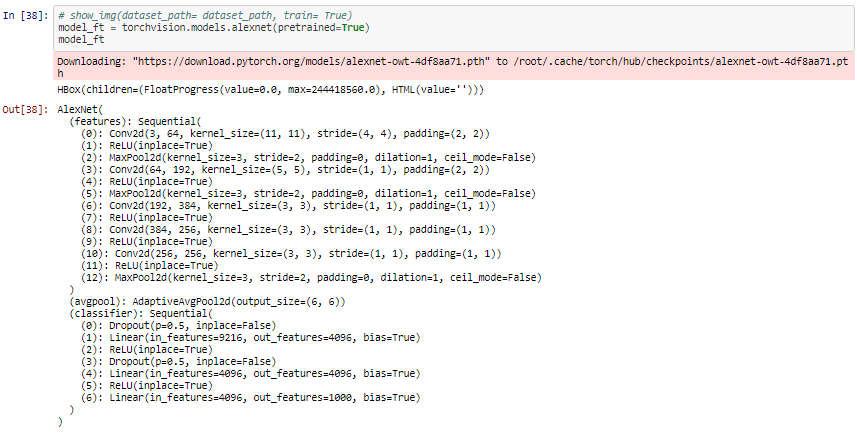


Figure 1.7 Alex-Net Architecture

Advantages:

* It is the first major CNN model that used GPU’s to train and this lead to faster training.
* It is a Deep Architecture with a total of 8 layers and this means efficient ability to extract features and it also works well with RGB images.
* It allows us to negate the negative outputs of summation of gradients and data and this means it’s speed increases much faster.
* Alex-Net uses “RELU” activation which means there won’t be much loss of features.
* “RELU” does not limit the output features which means not only dominant but also mild features are gained.

Disadvantages:

* Requires GPU for longer periods to train.

xix

Data-set Description:

In the view of In order to evaluate the performance of our feature extracting and classifying approach the Tawsifur Rehman’s covid-19 radiography database is used in our system which consists of 3616-COVID-19 positive along with 10,192-Normal, 6012- Lung Opacity (Non-COVID lung infection), and 1345-Viral Pneumonia images.

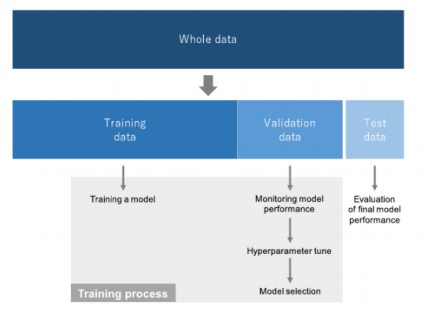


Figure 1.8 Data-set Split Diagram

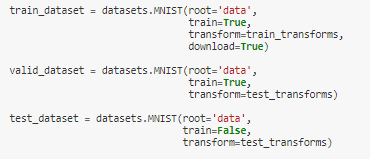


Figure 1.9 Data split Code

xx

Data Preprocessing:

Resizing: It means the default images are re-scaled to the size of 224x224

Greyscale to RGB: The Grad-Cam technique needs the data to be in the rgb type instead of greyscale.

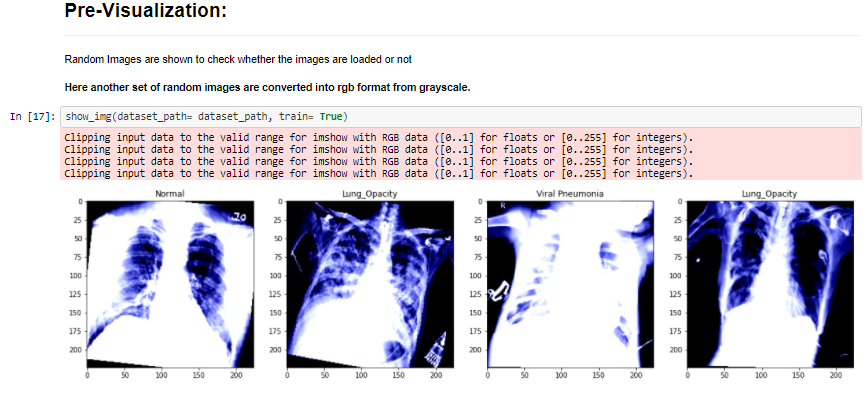


Figure 1.10 Random Images after RGB.

Performance Metrics:

A set of 4 performances criteria of deep transfer learning models those are:

Accuracy: This shows the no.of correct detected classes divided by the total number of test images.

Recall:Is the measure of covid-19 images that are correctly detected. Recall is very important in medical field.

Precision:Is the percentage of correctly detected labels in true positive patients.

F1- Score:is defined as the weighted average of precision and recall that combines both the precision and recall.

xx

5.Results & Discussion

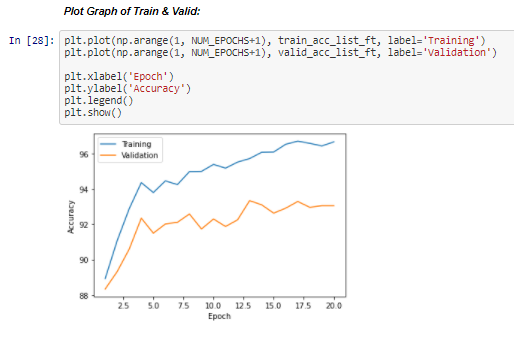


Figure 1.11 Model Accuracy Graph

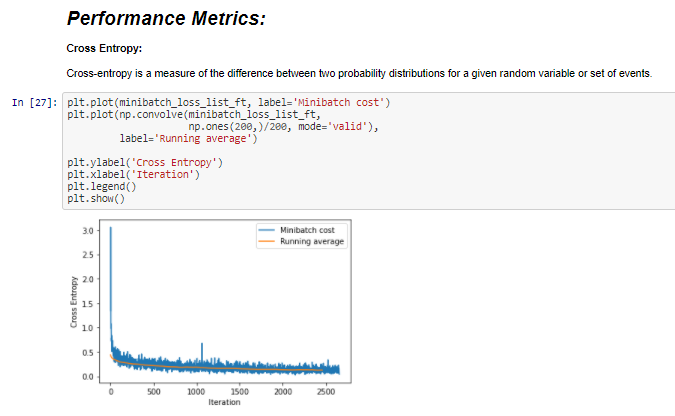


Figure 1.12 Model Cross Entropy Graph

xxi

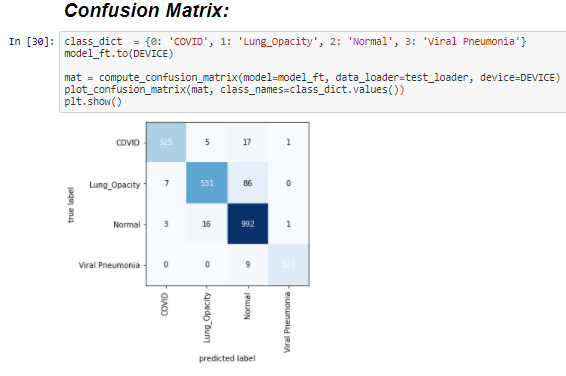


Figure 1.13 Confusion Matrix

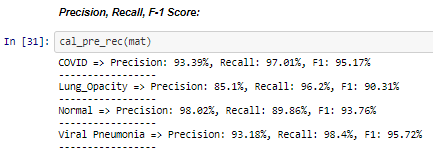


Figure 1.14 Precision, Recall, F-1 Scores of Each Class

xxii

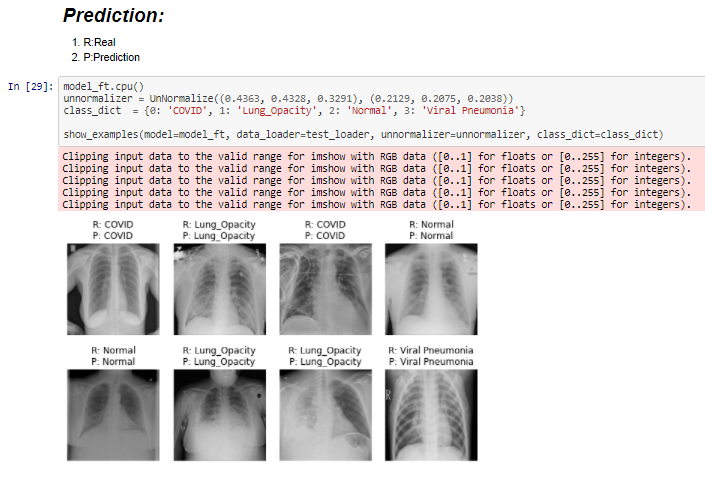


Figure 1.15 Model Prediction

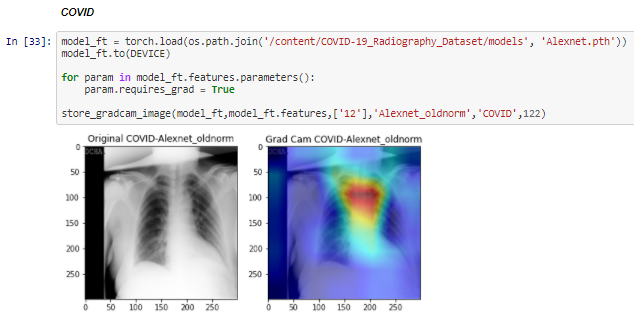


Figure 1.16 Grad-Cam For COVID

xxiii

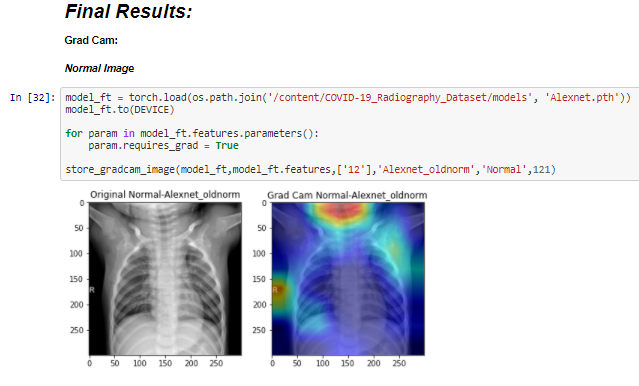


Figure 1.17 Grad-Cam for Normal

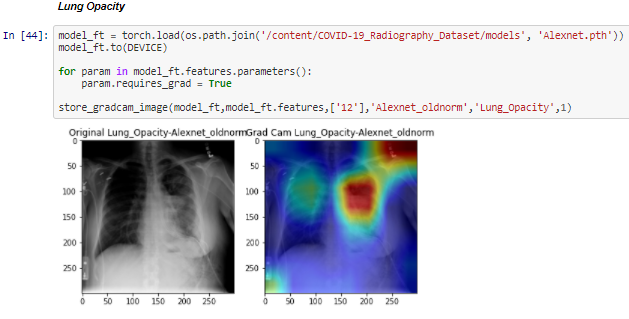


Figure 1.18 Grad-Cam for Lung Opacity

xxiv

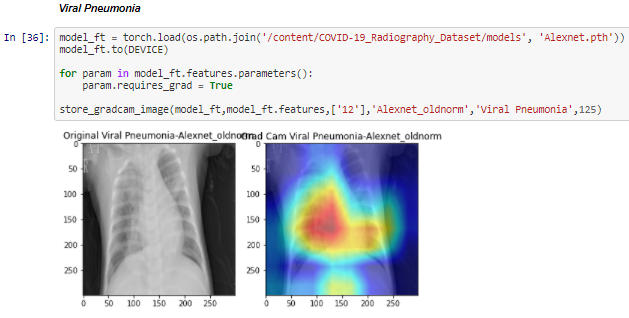


Figure 1.19 Grad-Cam for Viral Pneumonia

1. Conclusion & Future Work

The proposed system works ‘Efficiently and Accurately’ as compared to sequential convolutional network in detecting the presence of not just covid-19 but also some other abnormalities in the chest x-ray image and provides a way to visualize them using grad-cam technique and future work includes re-implementation using updated images and deployment of an application that allows medical personnel to use the system.

xxv

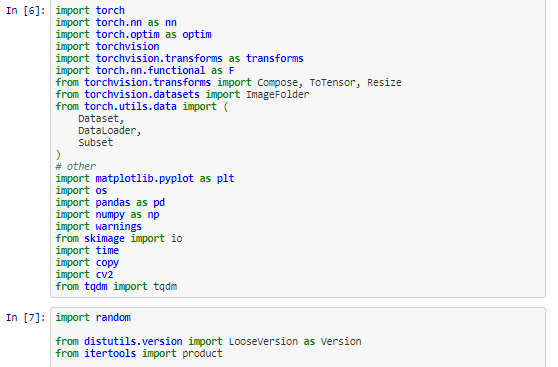
1. REFERENCES
2. M.E.H. Chowdhury, T. Rahman, “Can AI help in screening Viral and COVID-19 pneumonia”,Volume-8, Page(s): 132665 - 132676, 2020.
3. Tawsifur Rahmana, Amith Khandakara, “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images”, Volume-132,Pages(s):104319, 2021.
4. Qiblawey Y, Khurshid U, “COVID-19 infection localization and severity grading from chest X-ray images”, Volume-7, Page(s):12596-12598, 2020.
5. Mohammadi R 1, Salehi M, “Transfer Learning-Based Automatic Detection of Coronavirus Disease 2019 (COVID-19) from Chest X-ray Images”, Volume-5,page(s):559-568,2020.
6. Behera B1, Kumar N, “Covid-19 Detection Using Advanced CNN and X-rays”, Volume-348,Page(s):33-34, 2021.

Appendix:

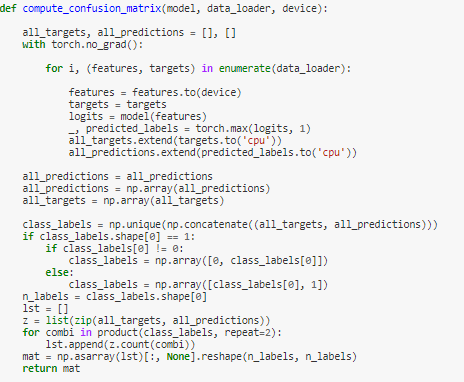
Dataset Forking:



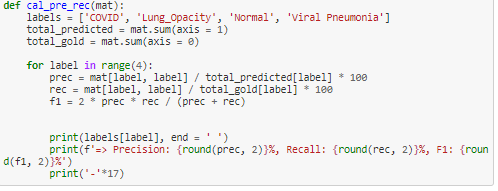
Imports:



Confusion Matrix:



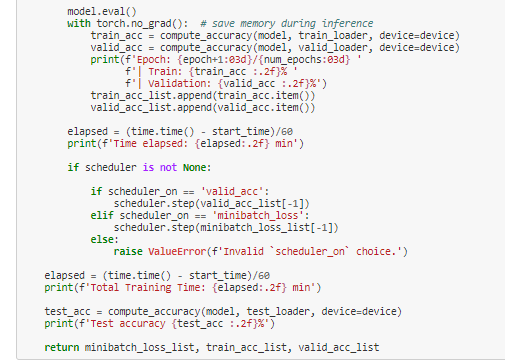
Precision, Recall, F1-score:

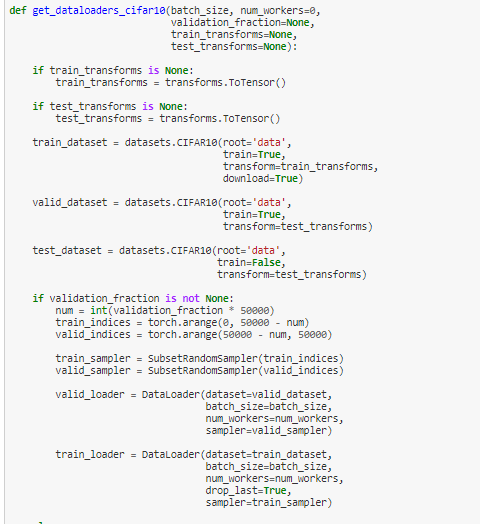


Training Model Layer Saving:

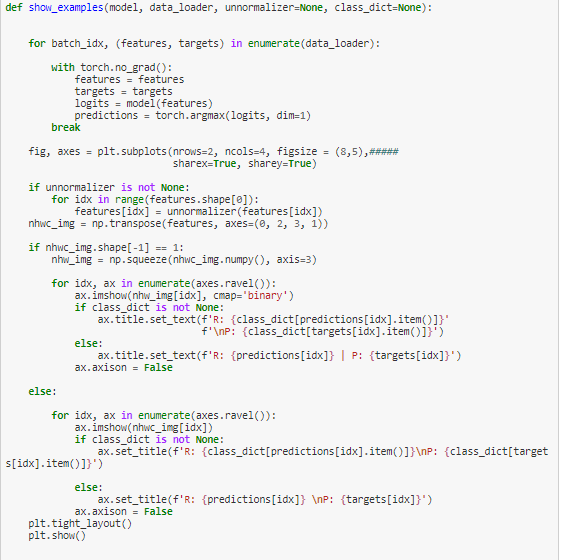


.contd:

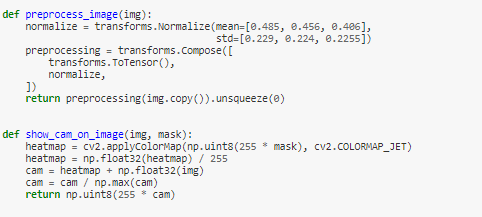




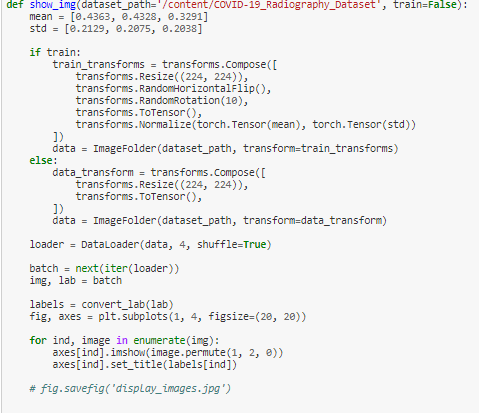
Prediction Function:



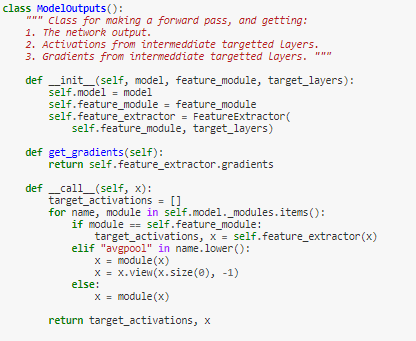
Pre-process Functions:



Greyscale to RGB



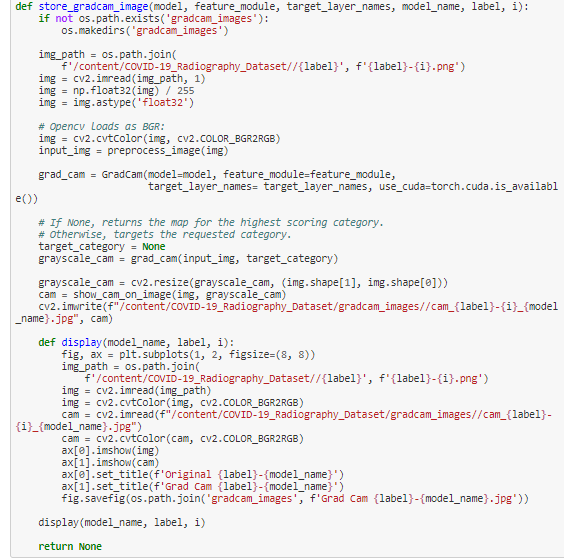
Model Modification:



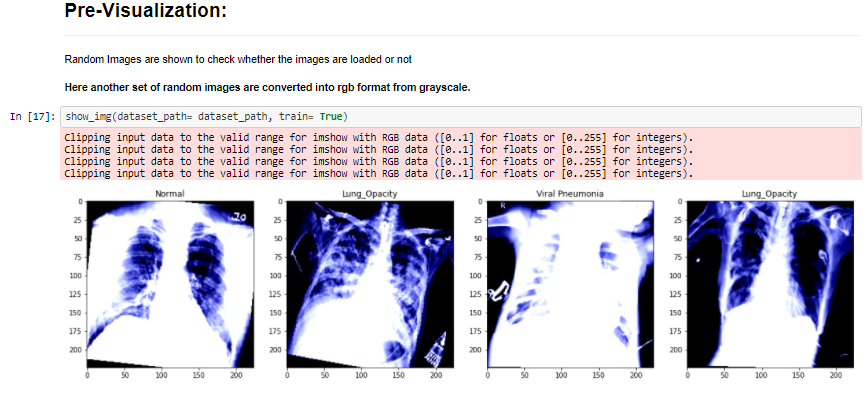
Model Parameters:



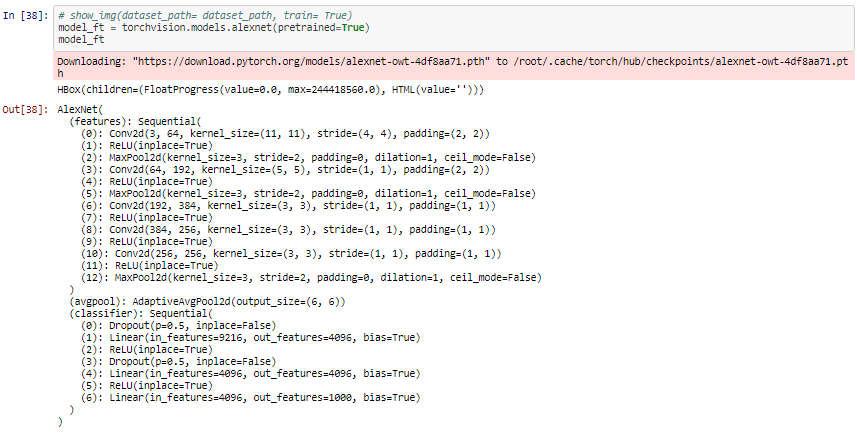
GRAD-CAM Function:



Pre-Visualization:



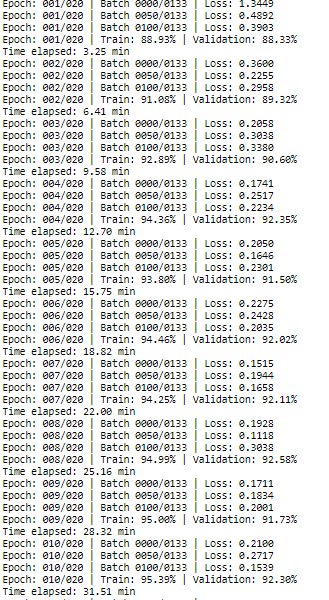
Model Loading:

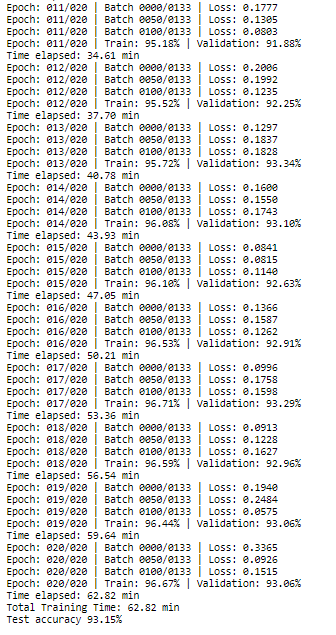


Model Building:

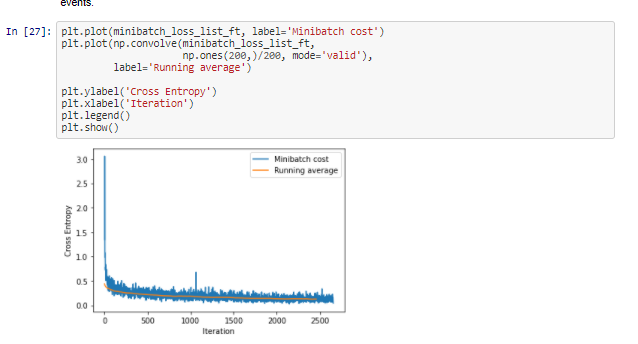


Training:

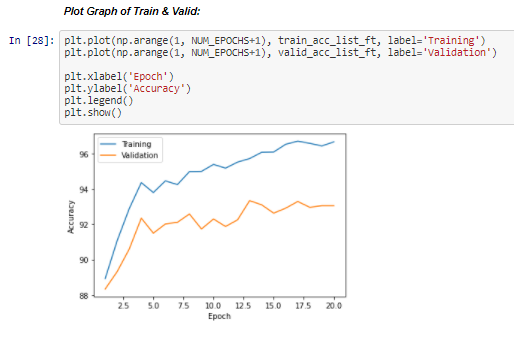




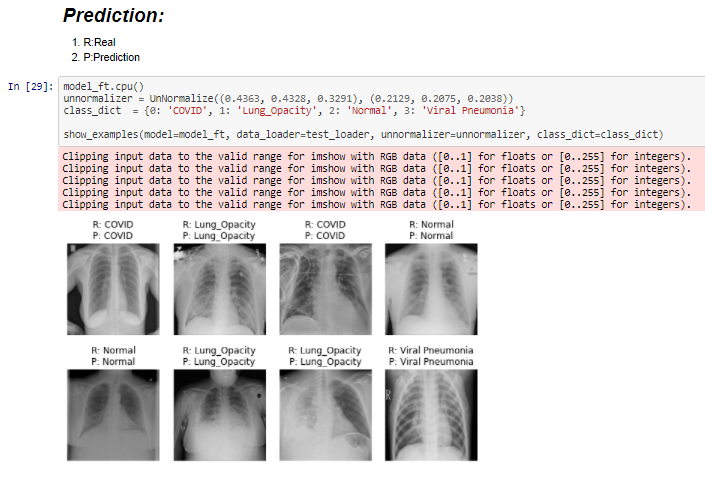
Cross Entropy:



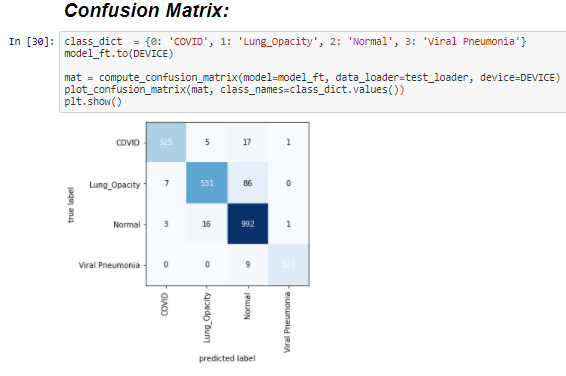
Model Train & Valid Accuracy:



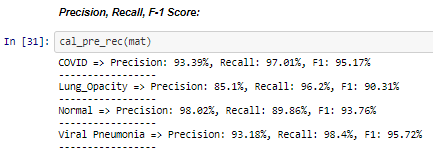
Prediction :



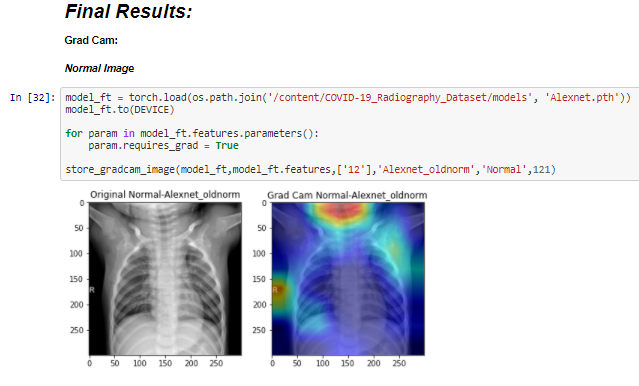
Confusion Matrix:



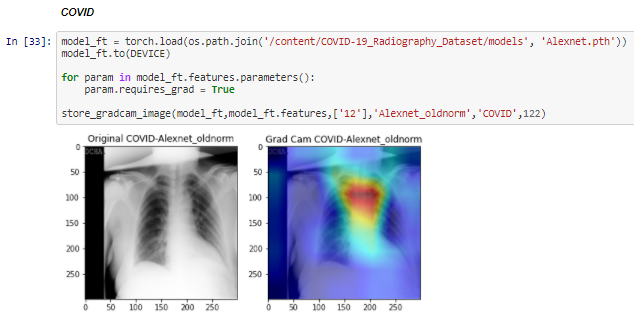
Precision Recall, F1-Score:



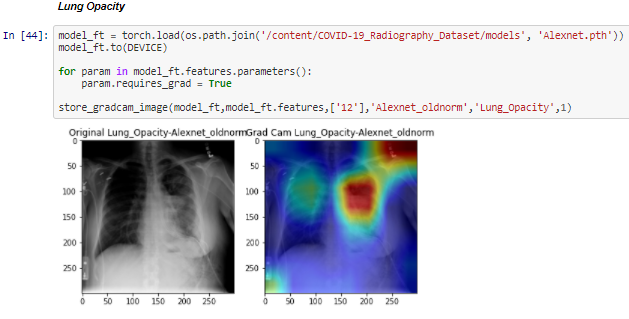
Grad-CAM for Normal:



Grad-Cam for COVID-19:



Grad-Cam for Lung Opacity:



Grad-Cam for Viral Pneumonia:

