COVID-19 PREDICTION USING LUNG X-RAYS

A report submitted under the partial fulfilment of the requirements for the degree of Bachelor of Technology in Computer Science and Engineering.

By,

L. NAGA SAI SRI RAVI TEJA – 121710307017

S. RITESH DEV – 121710307044

K. BHARATH – 121710307003

T. YASHWANTH – 121710307049

Under the guidance of

Dr. Don S. Kumar

Professor

GANDHI INSTITUTE OF TECHNOLOGY AND MANAGEMENT

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Deemed University declared under Section 3 of the UGC Act, 1956
Rushikonda, Visakhapatnam, Andhra Pradesh – 530045, India
www.gitam.edu

GANDHI INSTITUTE OF TECHNOLOGY AND MANAGEMENT

DEEMED TO BE UNIVERSITY

Deemed University declared under Section 3 of the UGC Act, 1956 (NAAC Acredited with A+ Grade)
Rushikonda, Visakhapatnam, Andhra Pradesh – 530045, India

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"COVID-19 PREDICTION USING LUNG X-RAYS"** carried out by **L. Naga Sai Sri Ravi Teja (121710307017), S. Ritesh Dev (121710307044), K. Bharath (121710307003), T. Yashwanth (121710307049)** in partial fulfilment of for the award of degree Bachelor of Technology in Computer Science and Engineering, GITAM – Deemed to be University, Visakhapatnam during the academic year 2017-2021.

(Signature)
Project Guide
Dr. Don S. Kumar
Professor

(Signature)
Head of the Department
Dr. Sireesha R.
Professor

DECLARATION

I hereby declare that the project entitled **"COVID-19 PREDICTION USING LUNG X-RAYS "** submitted to GITAM, Deemed to be University for the award of the degree of Batchelor of Technology in Computer Science and Engineering is a result of the original research carried out in this thesis. We understand that our report can be made electronically available to the public. It is further declared that the project report or any part there of has not been previously submitted to any University or Institute for the award of degree or diploma.

Name of Student(s): L. Naga Sai Sri Ravi Teja, S. Ritesh Dev, K. Bharath, T. Yashwanth
Admission No.s: 121710307017, 121710307044, 121710307003, 121710307049
Degree: Bachelor of Technology
Department: Computer Science and Engineering
Title of Project: COVID-19 DETECTION USING LUNG X-RAYS

Date: 06/04/21

ACKNOWLEDGEMENT

We take this opportunity to remember and acknowledge the cooperation, good will and support both moral and technical extended by several individuals out of which our project has evolved. We shall always cherish our association with them.

We are greatly thankful to Guide, **Dr. Don S. Kumar** for providing us with all his valuable suggestions, who has supported us throughout this project with his patience and guidance to make our project a success.

We would like to express thank and our gratitude to Head of the Department of Computer Science and Engineering, **Dr. Sireesha R.**, whose suggestions and encouragement have immensely helped us in the completion of the project and for their support and valuable suggestions during the dissertation work.

We offer our sincere gratitude to our Project Co-ordinator **Dr. Sireesha R.**, who has supported us throughout this project with her patience for her valuable guidance and encouragement during our project reviews.

We offer our sincere gratitude to our Project reviewer & A.M.C, **Shri. Bhargav K**, who has supported us throughout this project with his patience for his valuable guidance and encouragement during our project reviews.

We are extremely grateful our Parents and Friends for their blessings and prayers for our completion of project that gave us strength to do our project.

Submitted by,

L. Naga Sai Sri Ravi Teja – 121710307017

S. Ritesh Dev – 121710307044

K. Bharath – 121710307003

T. Yashwanth – 121710307049

ABSTRACT

The Covid-19 (also known as SARS-COV-2) that first occurred in Wuhan, 2019 which spread around the whole world like a wildfire. This contagious disease spreads from person to person through direct contact to another. The effects of Covid-19 can be classified into different scales from mild to severe. At the time of writing this paper a total of 148 million cases and 3.1 million deaths are confirmed. Most of the Covid-19 detection are done with RT-PCR tests which generally take time. Depending the critical scenarios and demands it might even take longer. For a contagious disease like covid-19 the main goal is to restrict it's spread. So, with the help of Machine Learning and Deep Learning Algorithms that are built on Radiology images could help in making the decisions for diagnosis of Covid-19 patients. We proposed in using Transfer-Learning based model for Covid-19 Detection using chest x-ray, because of the scarcity of available data. We performed Transfer Learning approach in order to obtain reliable results which could help us with smaller dataset. Though the x-rays do not provide maximum confirmation we rely the minimal percentage of chance that could help in reducing the spread of Covid-19. The process consists of two phases where in the first we pre-process the images and in the second we train and finetune the model to achieve desirable accuracy of the model.

Publicly available X-ray images (1583 healthy and 712 confirmed COVID-19) AND (712 COVID-19, 4273 Pneumonia and 1583 Normal) were used in the experiments, which involved the training of deep learning and machine learning classifiers. 5 custom CNN (Convolution Neural Network) experiments were, and 5 experiments for both categorical and binary were performed using Transfer Learning Models with ImageNet set as weights.

Keywords: Covid-19, SARS-COV-2, Deep Learning, Pre-Processing, Transfer Learning, Pandemic

TABLE OF CONTENTS

FIGURES

CHAPTER-1
INTRODUCTION

1.1. Motivation :

At the end of 2019, mankind faced a pandemic of Covid-19, since its emergence in Wuhan, China in December of 2019 [2]. The death toll was massive due to the unavailability of data on the cure for the disease. Even with restrictions of travel, the virus spread out throughout the world and the world has gone through a lockdown to reduce the spread of Covid-19. The death toll and spread rate was alarming for which social distancing norms were proposed by WHO(World Health Organization) [3]. Though some screening measures were taken, the spread of the virus did not go down and it reached a huge spike in various countries. The doctors were insufficient of the equipment they had and were affected by the virus losing lives. The treatment was given to the patients in an isolated environment with the necessary precautions [4]. Reverse-transcription polymerase chain reaction (RT-PCR) testing, which can detect SARS-COV-2 RNA is used to detect Covid-19 using respiratory specimens (such as nasopharyngeal and oropharyngeal swabs), is the best screening method for Covid-19 for clear confirmation of Covid-19. Though it provides accurate results the number of test kits available overshadowed the usage. Also with the amount of time taken to obtain the result is from (one day - or two days) [5]. Based on the demand sometimes it also takes a week. This in turn creates uncertainty in people and also increases the spread before the time we get the result. It is found in the earlier study that patients affected by covid-19 show abnormalities in the chest radiographs. So there is an immediate need of using radiology images for testing such as HRCT Scans and Chest X-Ray scans which could provide a visual indication of viral infection. Since Covid-19 affects the lungs, causing pneumonia. Finding out the abnormalities is easier and faster. The HRCT Scans are supposed to provide better image feedback and confirm whether a person has been affected by Covid or not [6]. As the facilities for chest imaging are readily available and also HRCT scans providing high accuracy it could act as an alternative to the RT-PCR tests. Through the X-Rays provide image feedback despite being subtle it is hard to confirm whether or not the person is affected by covid or not. So, the help of computer-aided deep learning or machine learning algorithms will make it more efficient and less time-consuming. And with the spike in the number of cases and cost of HRCT scans are too high to be available easily. So the X-Rays despite being cheaper it's faster to obtain and with help of Deep Learning models, we can use them to predict Covid-19 spread in a person. The goal of this study is to detect the chance of a person to have covid-19 thereby increasing the chance to reduce the spread of the virus and providing accurate treatment to the affected. This further helps in filtering out Covid-19 affected people.

The techniques of Deep learning always played a prominent role in making machines and software. Various terminologies like Max-Pooling, Convolutional Neural Networks (CNN), Sigmoid Function, and ReLU have prevailed in the world of Artificial Intelligence(AI). With the help of such technology, we can depict the disease.

The concept of CNN works fascinatingly. This is to mimic the neurons in the human brain. The only difference is that CNN used a methodology of updating the weights of the layers. A model of trained neural networks gains an understanding of the assigned knowledge. The main objective of AI is to mimic human behavior and intelligence. Deep Learning is capable of Image Recognition in which different models can be used to train with different architectures.

Radiological imaging such as X-rays and CT Scans can be used to apply advances AI Techniques and can be helpful in the prevention of the disease, and can also overcome the lack of physicians.

The proposed model is developed to provide diagnostics for binary classification (COVID vs. No-Findings) and Categorical classification (COVID vs Pneumonia vs Normal).

1.2 Project Objective

The main objective of our project is to prevent the spread of covid-19 using chest X-rays. The use of X-rays is due to its cost-effectiveness and is easily obtained in many remote villages. Though we do not propose the use of X-rays as a complete alternative to HRCT or RT-PCR tests these could help us in filtering out people and providing early treatment. On further thought, with the help of this method, we can perform selective tests of RT-PCR for restricting the spread and provide early treatment.

1.3 Advantages

1.3.1 Early detection and diagnosis of the infection:

ML can quickly analyze irregular symptoms and other 'red flags' and thus alarm the patients and the healthcare authorities. It helps to provide faster decision-making, which is cost-effective. It helps to develop a new diagnosis and management system for the COVID 19 cases, through useful algorithms.

1.3.2 Prevention of Spread:

The spread of covid-19 can be restricted and selective tests of RT-PCR can be done as the lesser availability of test kits.

1.4 Drawbacks of the Existing system

Though the currently existing system of testing with RT-PCR, Rapid Antigen, True Nat using Throat and Nasal Swabs provide accurate results, they take a considerable amount of time to provide the results (1day to few weeks) and labs are not readily available. The immediate unavailability of treatment causes the increase of virus spread.

1.5 X-Rays and CT Scans

1.5.1 X-Rays

Radiation that can pass through the body is known as X-rays. They aren't visible to the naked eye, and they aren't felt. The energy from X-rays is absorbed in various rates by different parts of the body as they move through it. After the X-rays have passed through, a detector on the other side of the body picks them up and converts them to an image.

Dense parts of your body that X-rays find it more difficult to move through, such as bone, show up as clear white areas on the picture. Darker areas indicate softer sections that X-rays can move through more quickly, such as the heart and lungs.

Figure 1: Shades of Grey

So, as seen in the picture, total black is on one end of the continuum, complete white is on the other, and in between, there are lesser blacks and lesser whites.

In X-rays:

1. It generates a two-dimensional image
2. It is mainly used to examine bones as well as to diagnose cancer and pneumonia.
3. The photographs are created using radiation.
4. It is the most used and widely accessible.
5. As opposed to other scans, it is cost-effective.

There are four to five densities in any given X-ray, such as:

Black: Air

Less black or greyish: Fat

Less white: They are soft tissues like heart/muscle or blood/fluid

White: It can be metal or calcium in bones

Very Bright: It can be a metallic object

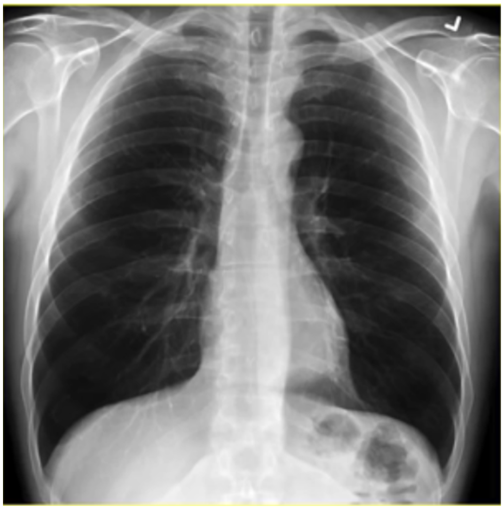


Figure 2: Lung X-Ray

And this is the sample image to better understanding

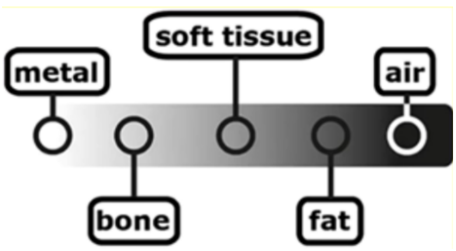


Figure 3: Shades from X-Ray

1.5.2 CT Scans

CT scans provide an accurate and high-quality image of the body. From different angles, processed variations of machine X-Rays are taken. Internal organs are photographed in 360 degrees. Doctors can diagnose medical conditions using the thorough quality of photographs of blood vessels, muscles, and other structures.

In CT scan

1. Create three-dimensional photographs.
2. It's also used to figure out what's wrong with organs and soft tissues.
3. It has a higher power level than X-rays.
4. The CT scan produces a 360-degree image.

Drawbacks of CT Scans:

1. It costs more.
2. CT scans are not available in every hospital.

1.6 Pneumonia in Lungs

Pneumonia is a lung infection that causes the air sacs of one or both lungs to become inflamed. Cough with phlegm or pus, fever, chills, and trouble breathing can occur when the air sacs fill with fluid or pus (purulent material). Pneumonia may be caused by several species, including bacteria, viruses, and fungi. [7].

1.6.1 Types of Pneumonia

1. Community-Acquired Pneumonia (CAP).
   - Community Acquired Pneumonia is when a patient gets it outside of the hospital.
2. Hospital-Acquired pneumonia (HAP).
   - Hospital Acquired pneumonia as the name suggests is when the patient gets it during their hospital stay.
3. Ventilator-Associated Pneumonia (VAP).
   - Ventilator Associated Pneumonia is when it’s acquired while the patient is on the mechanical ventilator.
4. Aspiration Pneumonia.

- Aspiration Pneumonia is acquired when a patient aspirates bacterium into the lungs, usually from food, saliva, or stomach acid.

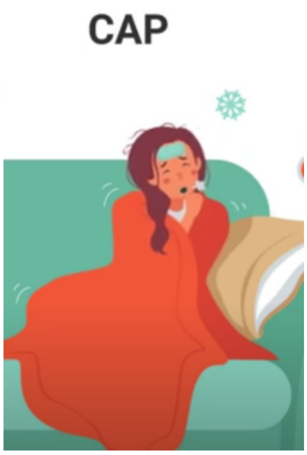


Figure 4: Community Aquired Pneumonia

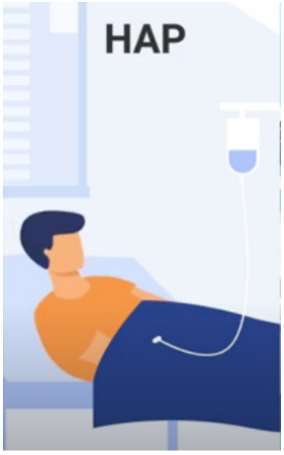


Figure 5: Hospital Aquired Pneumonia

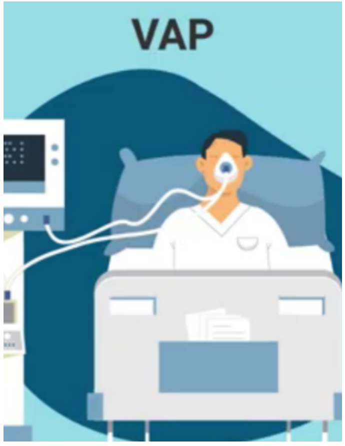


Figure 6: Ventilator Associated Pneumonia

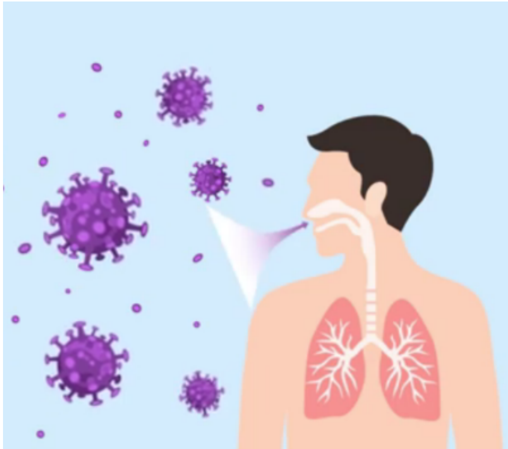


Figure 7: Aspiration Pneumonia

1.7 Covid-19

A cluster of pneumonia cases was recorded in the Chinese city of Wuhan in December 2019. The disease was caused by a newly identified coronavirus, according to investigations. Some of the early cases had registered visiting or working in a portion of seafood and live animal market in Wuhan. Covid-19 was given to the disease after that.

The Coronavirus is a big virus family.

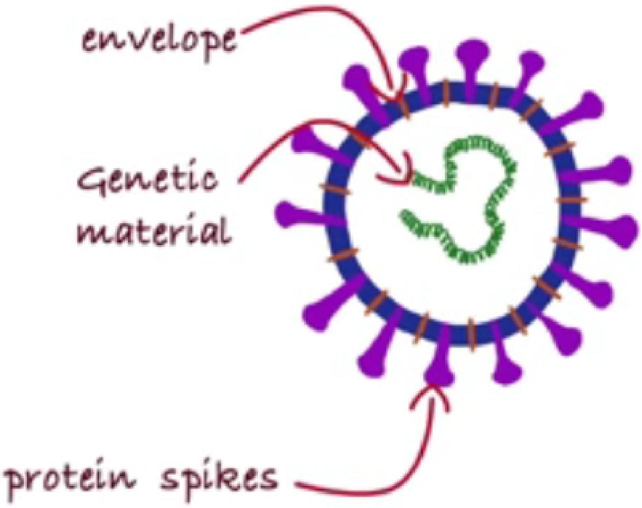


Figure 8: Covid-19 Strain

They are made up of a genetic material core encased in a lipid envelope with protein spikes. As a result, it has the appearance of a crown, as a crown is known or called Corona in Latin, and this is how these viruses got their names.

Pneumonia caused by Covid-19 is distinct from normal pneumonia in the following ways:

Covid-19 pneumonia symptoms may resemble those of other types of pneumonia or viral pneumonia. In the beginning, it was difficult for doctors and technicians to figure out which pneumonia they were dealing with because there were so many. As a result, it was difficult to guess or predict what was causing your illness.

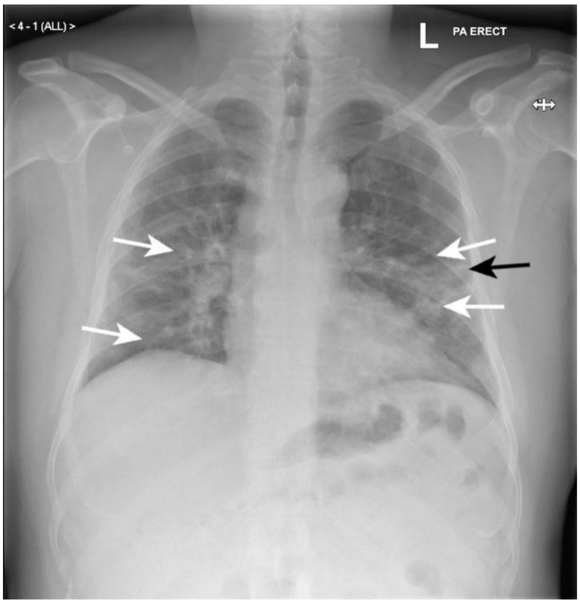


Figure 31: Covid-19 Pneumonia

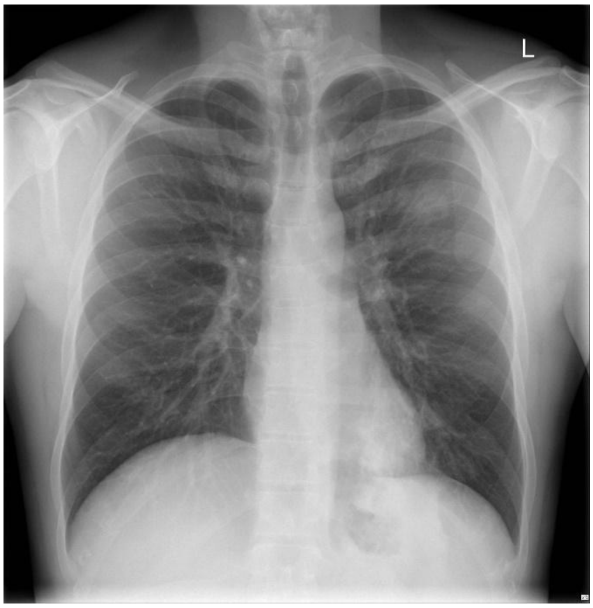


Figure 32: Non-Covid-19 Pneumonia

1.8 Convolutional Neural Network (CNN):

A convolutional neural network (CNN) is a neural network with one or more convolutional layers for image processing, classification, and segmentation. Convolution is a sliding filter that is applied to the data.

1.8.1 Benefits of CNN:

The use of pre-processing decreases the amount of human labour required to build its capabilities.
It's very easy to understand and put into practice.
Of all the algorithms that predict images, it has the highest accuracy.
Convolutional layers, pooling layers, and completely linked layers are the three types of layers that make up the CNN.

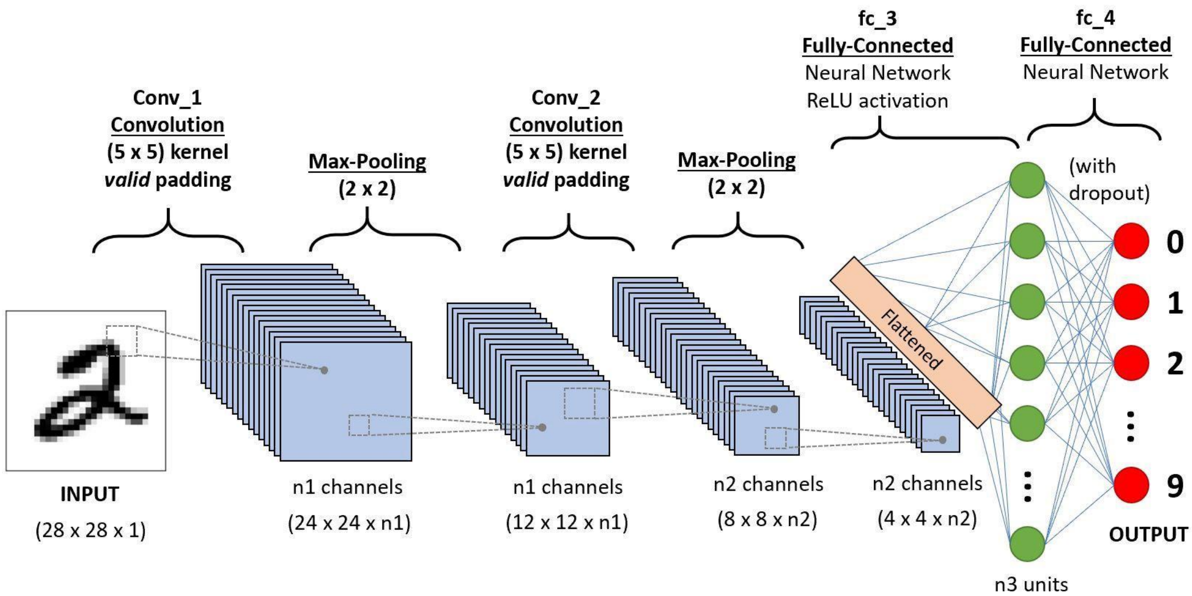


Figure 9: Convolution Neural Network

1.8.2 Input layer:

The input layer of a neural network is made up of artificial input neurons and is responsible for bringing the initial data into the system for processing by subsequent layers of artificial neurons. In an Artificial Neural Network, the workflow starts with an input layer.

1.8.3 Convolutional layers:

The input is convolved by convolutional layers, which then transfer the result to the next layer. This is similar to a neuron in the visual cortex reacting to a specific boost. Each convolutional neuron only calculates data for its sensitive area. While fully connected feedforward neural networks can be used to learn features and identify data, this architecture is unsuitable for larger data sets such as high-resolution images. Because of the enormous information size of images, where every pixel contains essential information, even a shallow design will necessitate a large number of neurons. For example, a completely associated layer for a (little) picture of size 100 x 100 has 10,000 loads for every neuron in the subsequent layer. Where all other factors are equal, convolution reduces the number of free boundaries, allowing for a more profound organisation. Using a 5 x 5 tiling locale, each with identical shared loads, for example, needs just 25 learnable boundaries, regardless of picture size. Using regularised loads over fewer boundaries avoids the disappearing slopes and detonating inclinations issues seen in traditional neural organisations during backpropagation. Furthermore, since spatial connections between independent highlights are considered during convolution and pooling, convolutional neural organisations are perfect for information with a grid-like topology (such as pictures).

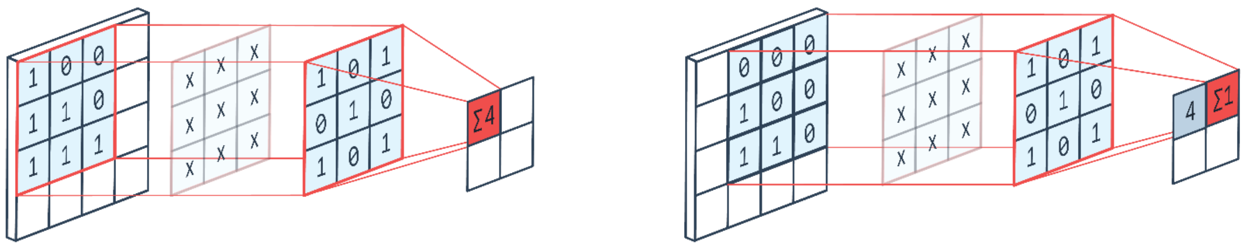


Figure 10: Process of Convolution

1.8.4 ReLU Activation Function:

The non-saturating activation function is described by ReLU, which is a truncation of the rectified linear unit. By setting negative qualities to zero, successfully removes them from an activation map. It introduces nonlinearities in the option capability as well as the overall network without affecting the convolution layers' open fields.

Different capacities, such as the saturating hyperbolic tangent and the sigmoid power, can also be used to create nonlinearity. ReLU is common among various capacities.

because it prepares the neural organisation many times faster without sacrificing generalisation accuracy.

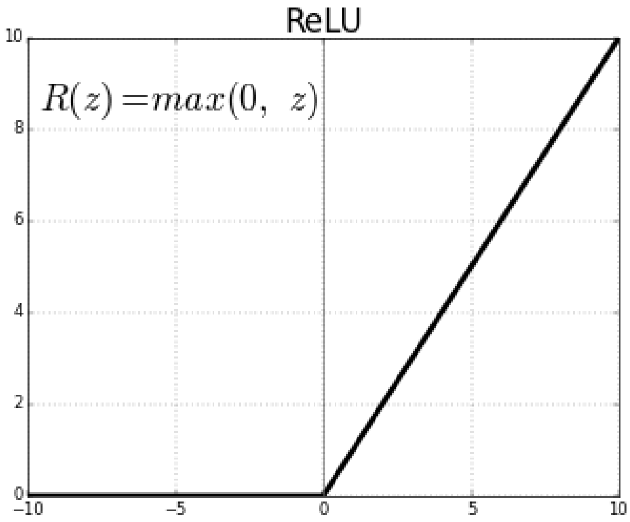


Figure 11: Relu Activation Function

1.8.5 Pooling Layers:

Along with traditional layers, convolutional networks can include local and/or global pooling layers. By joining the outputs of neuron groups at one layer into a solitary neuron in the next layer, pooling layers reduce the components of information. Local pooling connects small clusters; for example, tiling sizes of 2 x 2 are commonly used. Each of the component map's neurons is affected by global pooling. Two types of pooling are widely used: maximum and average. The maximum value of each local cluster of neurons in the feature map is used in peak pooling, while the average value is used in average pooling.

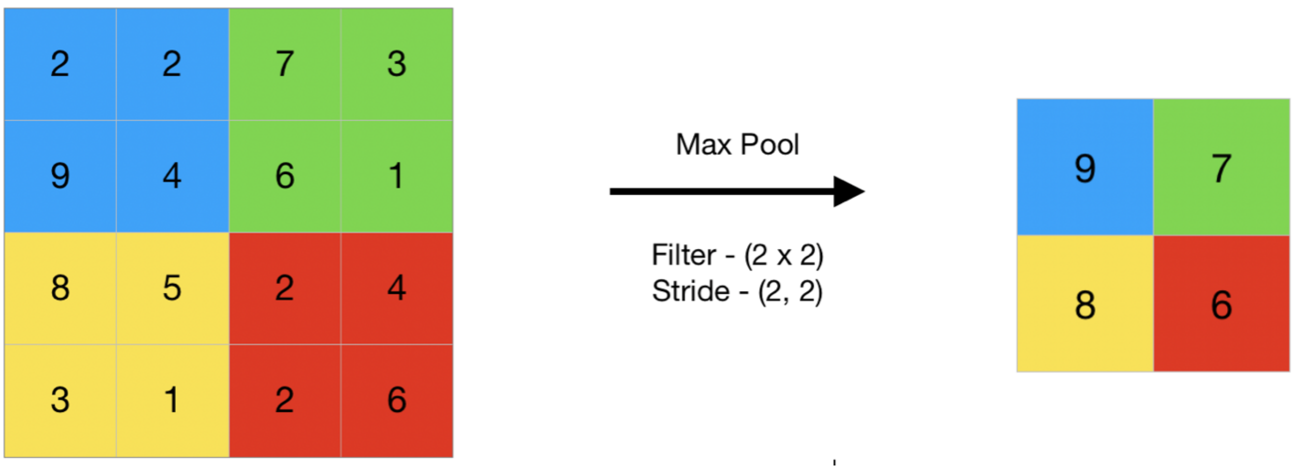


Figure 12: Max Pooling Example

1.8.6 Fully Connected Layers:

Every neuron in one layer is linked to every neuron in another layer in fully connected layers. It functions similarly to a multi-layer perceptron neural network. To identify the images, the flattened matrix tests a completely connected layer.

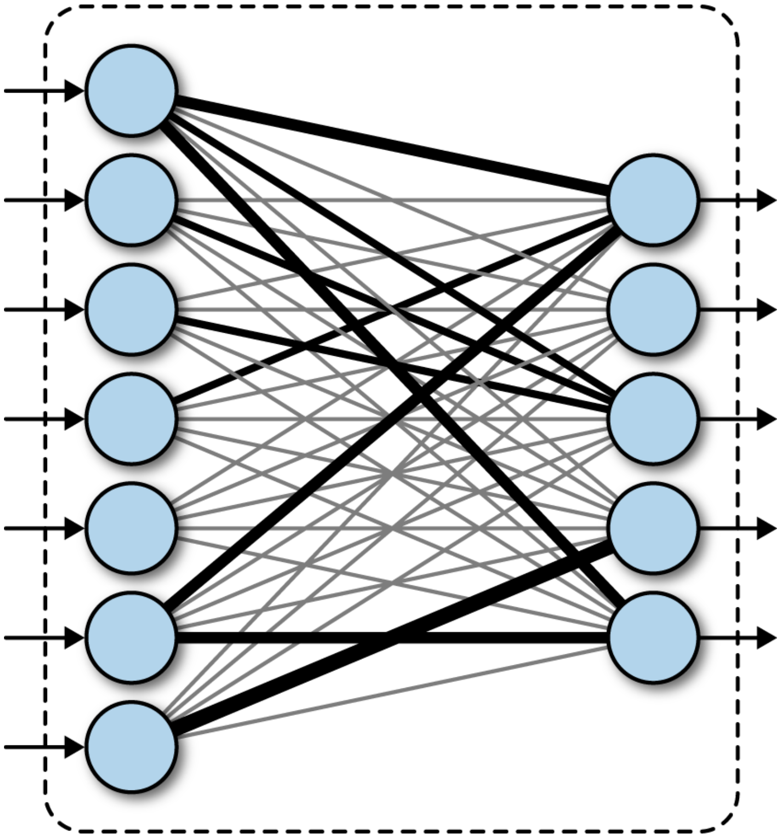


Figure 13: Fully Connected Layer

1.8.7 Weights:

Each neuron in a neural network registers an output value by applying a specific function to the input values obtained from the previous layer's receptive field. A vector of loads and a bias determine the function that is applied to the input values. Iteratively shifting these biases and loads is what learning entails.

Filters are defined as a vector of loads and a bias that address specific features of the input (e.g., a specific shape). The fact that several neurons can share the same filter is a distinguishing feature of CNN's. This decreases the memory impression because a single bias and a single vector of weights are used across all receptive fields that share that filter, instead of each open field having its own bias and vector weighting.

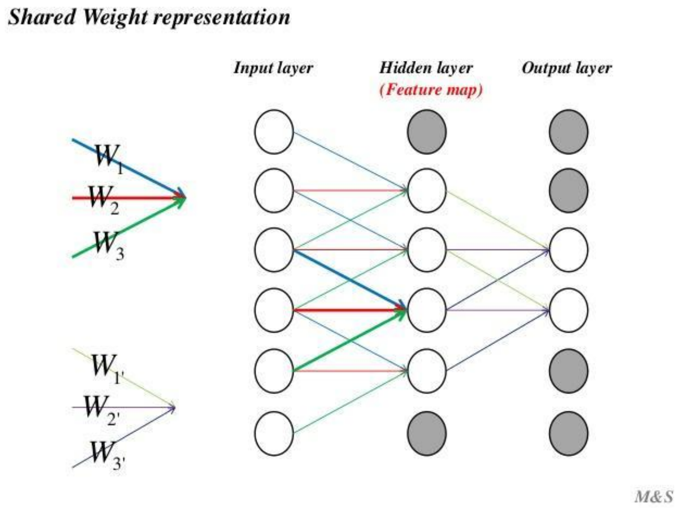


Figure 14: Weight Representation in CNN

1.8.8 Flatten Layer:

Flattening is the process of transforming data into a one-dimensional array that can then be input to the next layer. To build a single long feature vector, we flatten the output of the convolutional layers. It must also be linked to the final classification model, which is referred to as a fully connected layer.

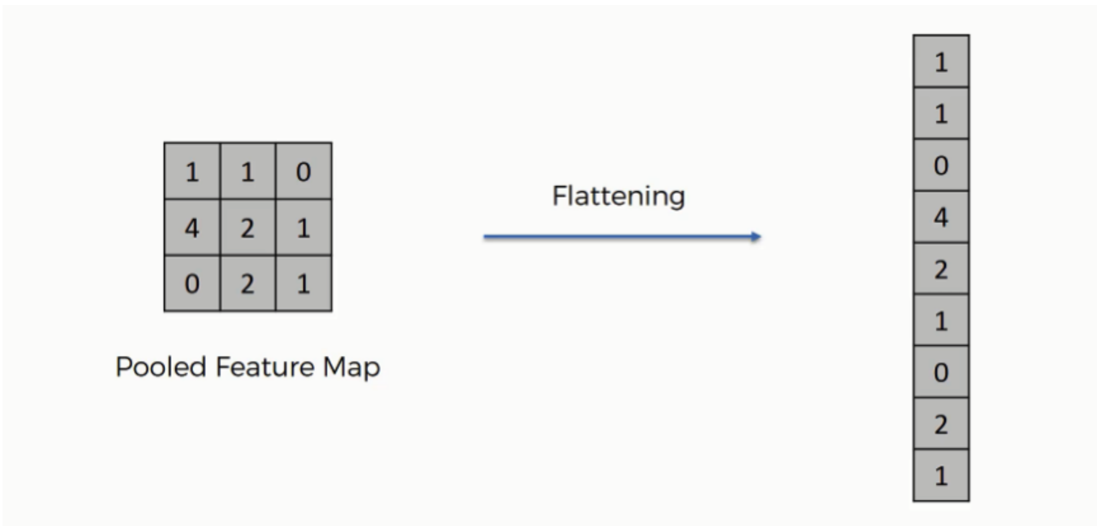
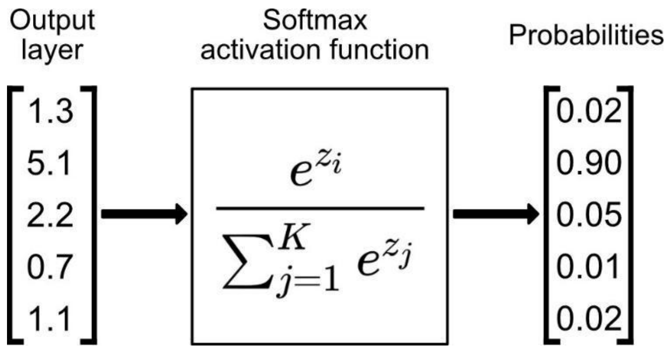


Figure 15: Flattening in CNN

1.8.9 SoftMax Layer:

The SoftMax function converts a vector of K real values into a vector of K real values that add up to one. As a result, it's common to add a SoftMax feature as the neural network's final layer.



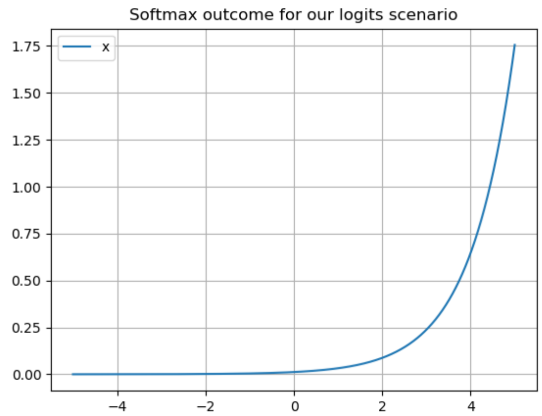


Figure 16: SoftMax Function

Figure 17: SoftMax Graph

1.8.10 Sigmoid Layer:

We use the sigmoid function since it occurs between two points (0 to 1). As a result, it is particularly useful in models where the likelihood must be predicted as an output.

Since the chance of something only exists between 0 and 1, sigmoid is the best option.

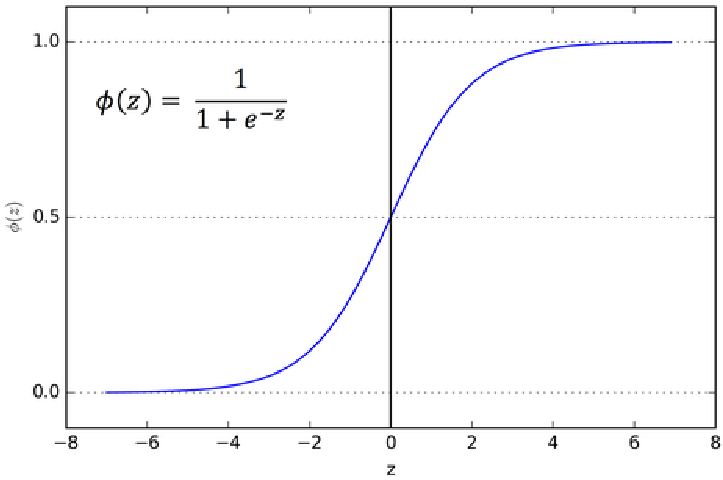


Figure 18: Sigmoid Activation Function

CHAPTER-2

LITERATURE SURVEY

2.1. Novel Feature Selection and Voting Classifier Algorithms for COVID-19 Classification in CT Images by IEEE – Published on IEEE Access

Using two datasets, this article proposes two optimization algorithms for COVID-19 feature selection and classification (Covid-19 and Non-Covid-19).

The paper is broken down into three parts. Using a Convolutional Neural Network (CNN) called Alex Net, the parameters (fitness and selected size) are primarily taken into account from CT scans. Second, a proposed feature selection algorithm based on Stochastic Fractal Search (SFS), Directed Whale Optimization Algorithm (Guided WOA), is used to improve the classification performance. Finally, a proposed voting classifier, Directed WOA based on Particle Swarm Optimization (PSO), considers the predictions of different classifiers to choose the most common class. And three separate terminologies are responsible for the result. [8]

2.2. Deep learning-based detection and analysis of COVID-19 on chest X-ray images

This research paper gave us details on how to use X-rays for Covid-19. The given paper developed models based on state-of-the-art CNN architectures, which were then modified to meet the model's requirements. According to the paper, the XCeptionNet has the highest output and is the best fit for use of the three models (Inception V3, Xception, and ResNet). However, the information we needed was for determining the dependability of X-Rays, which we were able to obtain with the aid of this paper. [9]

2.3. Identifying COVID19 from Chest CT Images - A Deep Convolutional Neural Networks Approach

The use of Ensemble learning in producing the most reliable results is proposed in this paper. Individual baseline models are first thoroughly tested in this phase. VGG16, InceptionV3, ResNet50, DenseNet121, and DenseNet201 are among the baseline versions. The convolution sections of both of these baseline models are kept the same as the regular models, as proposed for the ImageNet challenge; however, the fully connected parts of the models are set as three fully connected layers (4096, 4096, and 1000), each with ReLU activation, and a single-node prediction layer with Sigmoid activation feature. [10]

2.4. XCOVNet: Chest X-ray Image Classification for COVID-19 Early Detection Using Convolutional Neural Networks

The researchers created a convolutional neural network (CNN) to classify patients' chest X-ray images as positive COVID+ or negative COVID. A CNN with the Adam optimizer and a learning rate of 0.001 is used in our XCOVNet model. It uses a handcrafted seed dataset for CNN local and global features with 196 COVID+ patient chest X-ray images and 196 COVID images and does not include any feature selection process.

The proposed XCOVNet model, which is based on computerised automatic detection, is more effective at understanding features and detecting COVID19 than other traditional learning methods. The trained CNN comprises three convolutional layers with the kernel size of 3 × 3 followed by a rectified linear unit (ReLU) activation function which takes input images of size 224 × 224 × 3. The proposed XCOVNet system achieved an accuracy of 98.44% in classifying chest X-ray images. [11]

CHAPTER 3
SYSTEM ANALYSIS

3.1 Hardware Requirements

1. A processor of 4.16 GHz least is required.
2. Needs a system with a least of 8 GB RAM.
3. 64-bit architecture.
4. A least storage of 500GB.
5. Nvidia 1080Ti GPU

3.2 Software Requirements

1. Windows Operating System.
2. Most modern chrome browser.
3. Web Development for UI.
4. Python for coding.
5. Flask for the Backend.

3.3 Feasibility Analysis

A feasibility analysis ensures that considerations such as the project's legality, technological feasibility, and economic justification are all met. Economic problems cannot always be resolved, and resource issues, i.e., resources can become scarce for the task at hand.

3.3.1 Economical Feasibility

The analysis of the cost/benefits of the project, by this we can estimate the cost/benefits of the project before allocating the budget. The Present project that we are doing is economically feasible because of using Python and Angular which are open source.

3.3.2 Technical Feasibility

The key focus is on ensuring that the technological resources match the capability of the organization's technical viability, which involves both software and hardware specifications for the project.

CHAPTER-4

SYSTEM DESIGN

4.1 System Architecture:

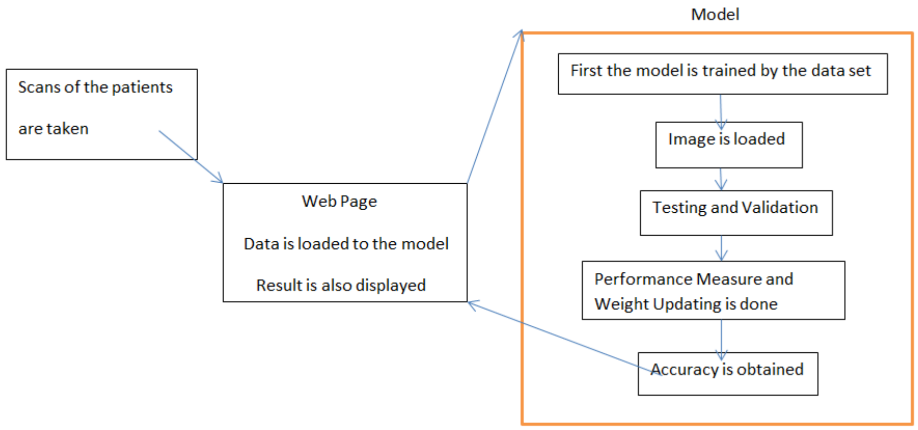


Figure 19: Architecture of the overall project.

4.2 Class Diagram:

A class diagram depicts the attributes (the knowledge they contain) and operations (the job they perform).

The structural diagram is another name for the class diagram.

The primary goal of creating the class diagram is to clarify the system's responsibilities.

The arrows indicate the relationship between the schools (Dependency, Association, and Generalization).

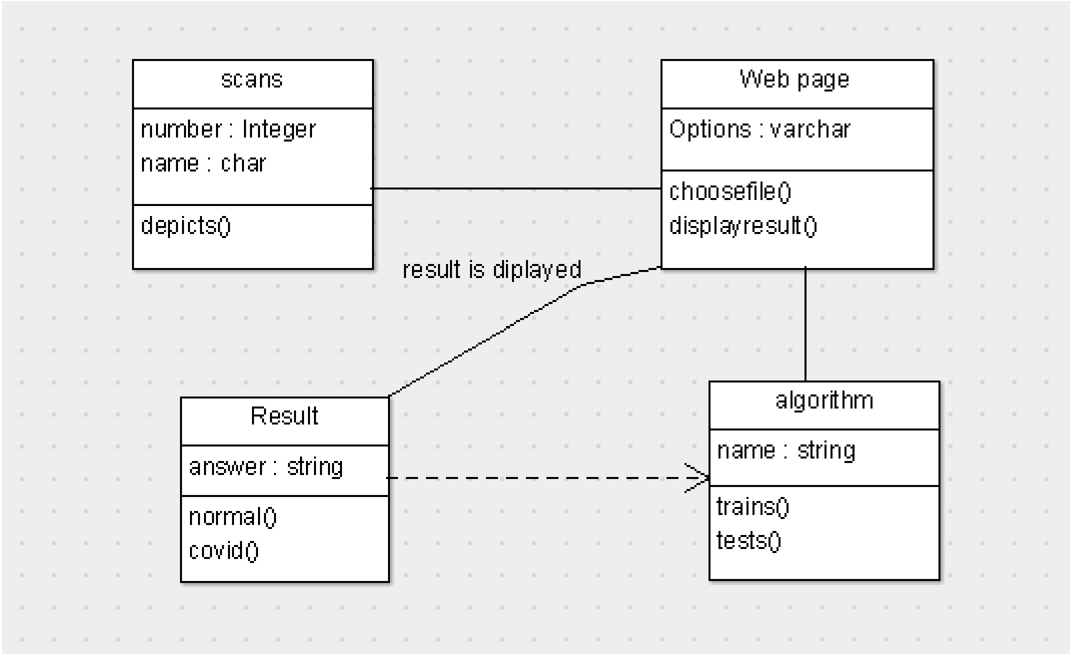


Figure 20: Class Diagram

4.3 Activity Diagram:

It's a flowchart that depicts the transition from one operation to the next. The task explains how the mechanism works.

The flow of control is from one process to the next.

The machine will be in the idle stage in the beginning because there will be no operation.

The first dot is the Initial state, which is where the process will begin.

The final state is where the whole process comes to a close.

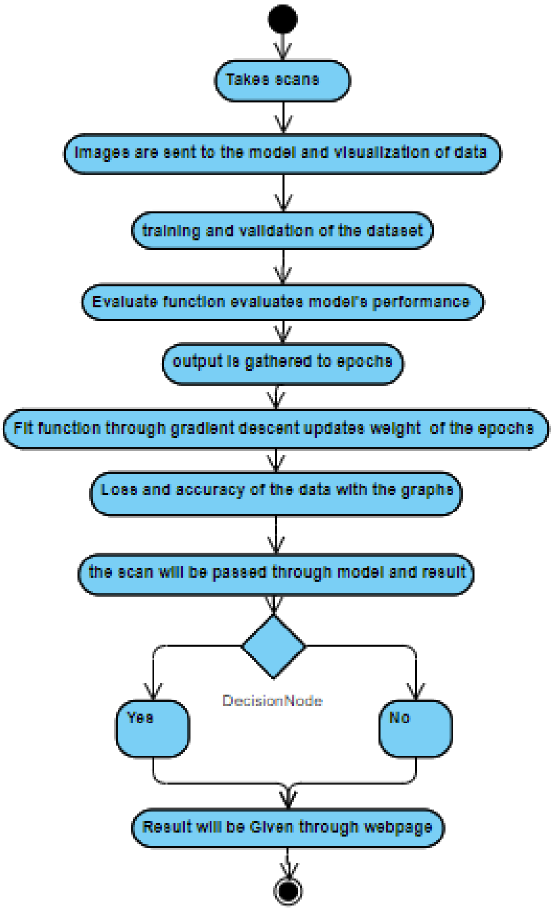


Figure 21: Activity Diagram

4.4 Use Case Diagram:

One of the complex natures of UML diagrams is the use case diagram. These diagrams are used to describe the functionality of the system's means and the tasks that it performs.

Actors, use cases, links, and the system's boundary are all found here.

Actors are the individuals that deal with the use cases or functions. They are the ones who started the use cases, and they have high expectations for the outcome.

Use cases are machine functions that are linked to actors; however, some use cases are linked to other use cases or may initiate their own processes.

Actors and the use cases are linked through a communication link, showing message passing between them.

Any use case beyond the system boundary is no considered because they will be not in the process.

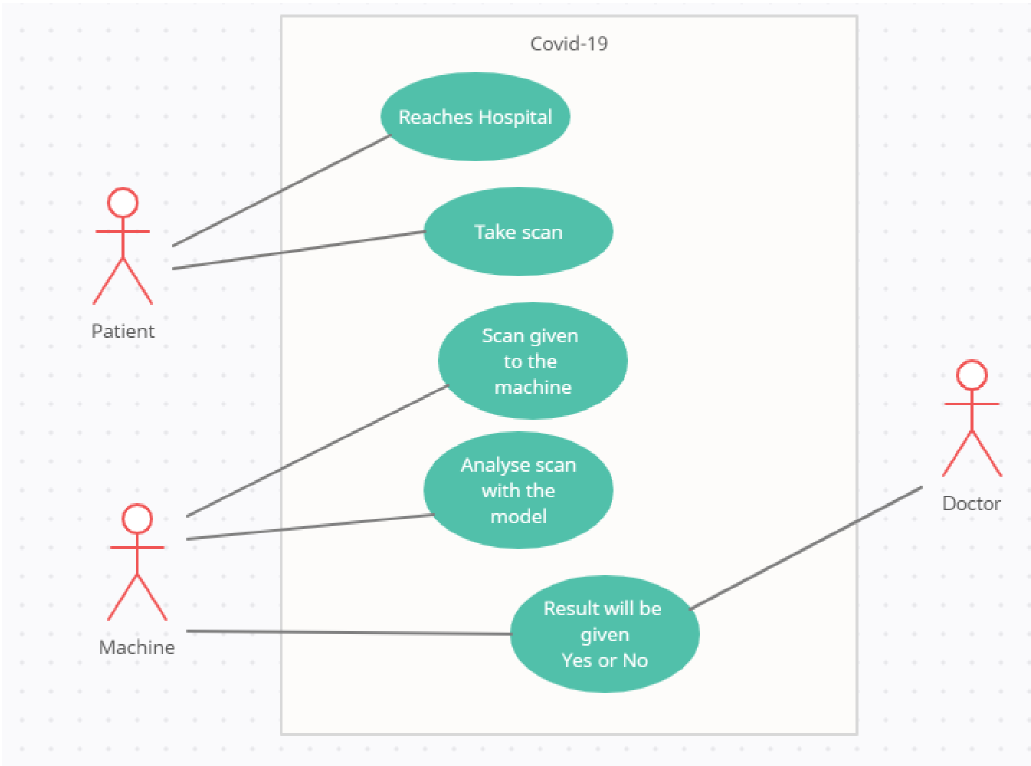


Figure 22: Use Case Diagram

4.5 Sequence Diagram

Sequence Diagrams are interaction diagrams in which the detailed description of the system's work process is illustrated.

The whole process is presented in the form of a number in sequential order so that everyone can understand what the machine is doing.

The objects are listed here, as well as the interactions between them so that we can learn about the important objects in the entire system.

The sequence diagram's key argument is that only the objects involved in the process are seen, with no other objects included.

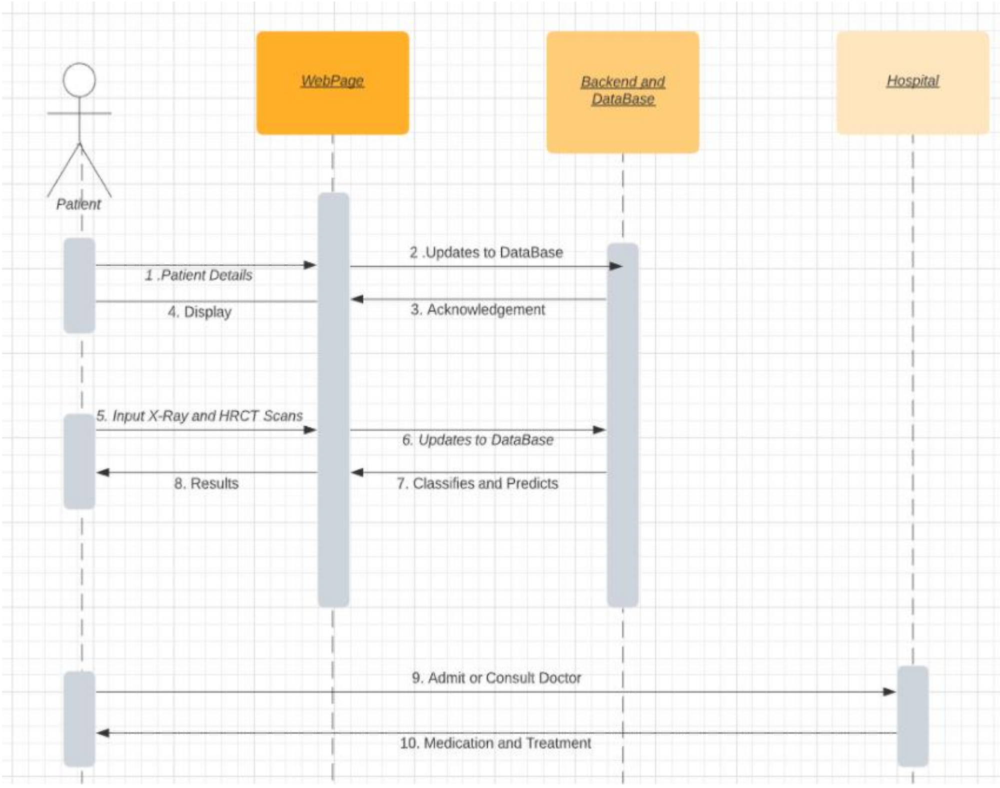
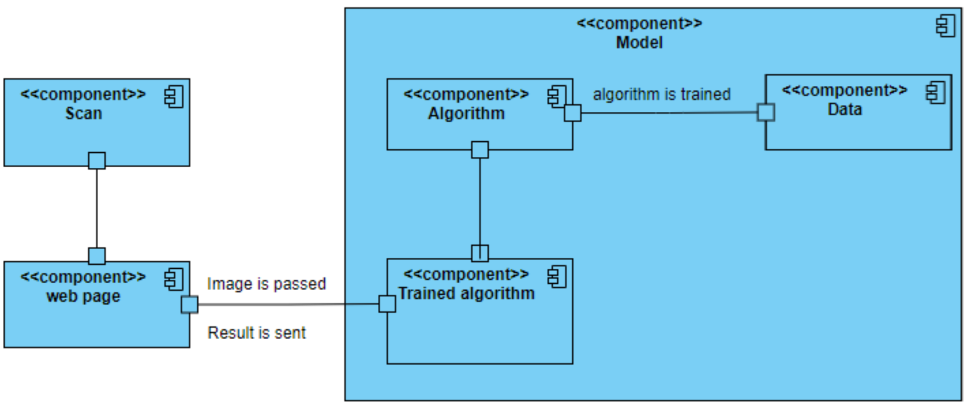


Figure 23: Sequence Diagram

4.6 Component Diagram

The physical objects of a system are depicted in component diagrams. The physical aspects of a node are the components such as libraries, archives, records, and so on. The data in this diagram will be represented by nodes in the form of files, records, codes,



and so on.

Figure 24: Component Diagram

4.7 Deployment Diagram

Deployment diagrams are used to show where the device modules are deployed on the hardware. These are used for the system's static view.

It has nodes and relationships, with nodes representing the elements.

This diagram depicts how the software component's data would be exchanged with the hardware components.

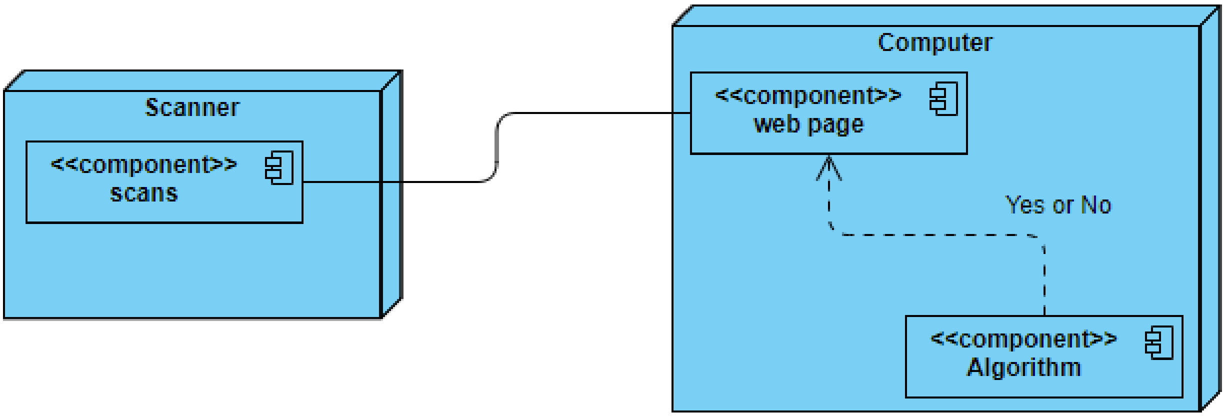


Figure 25: Deployment Diagram

CHAPTER-5

MATERIALS AND METHODOLOGY

5.1 Materials:

5.1.1 Image Data Set Collection:

The data for this study was gathered from the Kaggle repository [1.a], which contains normal and pneumonia chest X-ray scans, and the GitHub Repository of education454 [1.b], which contains normal and covid-19 chest X-ray scans.

The collected Dataset does not claim to improve the Deep Learning model's diagnosis; rather, it aids in the research of efficiently detecting Covid-19 infections using computer processing techniques. The total number of Lung X-Ray images in the dataset is 6568. ['covid', 'pneumonia', 'normal'] are the directories mentioned. The images seemed to be in-consistent, and the below table (Table1) provides the distribution of data. Since the images are equally distributed the classification can be more reliable.

5.2 Data Set Analysis:

Table 1: Dataset Contents

Covid -19	712
Pneumonia	4273
Normal	1583

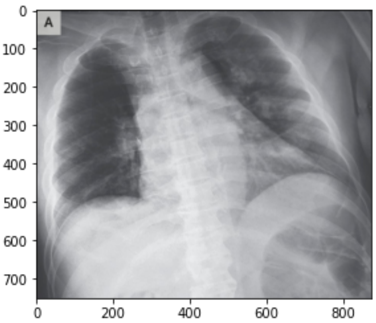


Figure 26: Sample Image

Table 2 & 3: Data Frame of Image File Set

The image in the dataset is classified in a data frame as below which is split into train and test data to understand the information we have.

	class	o_directory
0	PNEUMONIA	PNEUMONIA/person1493_bacteria_3896.jpeg
1	PNEUMONIA	PNEUMONIA/person54_bacteria_257.jpeg
2	PNEUMONIA	PNEUMONIA/person896_virus_1548.jpeg
3	PNEUMONIA	PNEUMONIA/person441_bacteria_1910.jpeg
4	PNEUMONIA	PNEUMONIA/person72_bacteria_352.jpeg
...	...	...
1630	NORMAL	NORMAL/NORMAL2-IM-1094-0001-0002.jpeg
1631	NORMAL	NORMAL/NORMAL2-IM-0452-0001.jpeg
1632	NORMAL	NORMAL/IM-0189-0001.jpeg
1633	NORMAL	NORMAL/NORMAL2-IM-0684-0001-0001.jpeg
1634	NORMAL	NORMAL/NORMAL2-IM-0583-0001.jpeg

1635 rows × 3 columns

	class	o_directory
0	PNEUMONIA	PNEUMONIA/person1952_bacteria_4883.jpeg
1	PNEUMONIA	PNEUMONIA/person1947_bacteria_4876.jpeg
2	PNEUMONIA	PNEUMONIA/person1949_bacteria_4880.jpeg
3	PNEUMONIA	PNEUMONIA/person1950_bacteria_4881.jpeg
4	PNEUMONIA	PNEUMONIA/person1954_bacteria_4886.jpeg
...	...	...
802	NORMAL	NORMAL/NORMAL2-IM-0354-0001.jpeg
803	NORMAL	NORMAL/NORMAL2-IM-0030-0001.jpeg
804	NORMAL	NORMAL/NORMAL2-IM-0369-0001.jpeg
805	NORMAL	NORMAL/NORMAL2-IM-0373-0001.jpeg
806	NORMAL	NORMAL/NORMAL2-IM-0271-0001.jpeg

807 rows × 3 columns

Table 4 & 5: Data Frame for Binary Classification

The image in the dataset is classified into binary dataset as the data frame as below for to have easier understanding of the information.

	class	o_directory
0	COVID19	COVID19/COVID19(415).jpg
1	COVID19	COVID19/COVID-19 (313).jpg
2	COVID19	COVID19/COVID19(63).jpg
3	COVID19	COVID19/COVID-19 (371).jpg
4	COVID19	COVID19/COVID19(512).jpg
...	...	...
479	NORMAL	NORMAL/NORMAL(561).jpg
480	NORMAL	NORMAL/NORMAL(1402).jpg
481	NORMAL	NORMAL/NORMAL(301).jpg
482	NORMAL	NORMAL/NORMAL(219).jpg
483	NORMAL	NORMAL/NORMAL(1416).jpg

484 rows × 3 columns

	class	o_directory
0	COVID19	COVID19/COVID19(445).jpg
1	COVID19	COVID19/COVID19(409).jpg
2	COVID19	COVID19/COVID19(373).jpg
3	COVID19	COVID19/COVID19(575).jpg
4	COVID19	COVID19/COVID19(497).jpg
...	...	...
1806	NORMAL	NORMAL/NORMAL(612).jpg
1807	NORMAL	NORMAL/NORMAL(871).jpg
1808	NORMAL	NORMAL/NORMAL(367).jpg
1809	NORMAL	NORMAL/NORMAL(90).jpg
1810	NORMAL	NORMAL/NORMAL(502).jpg

1811 rows × 3 columns

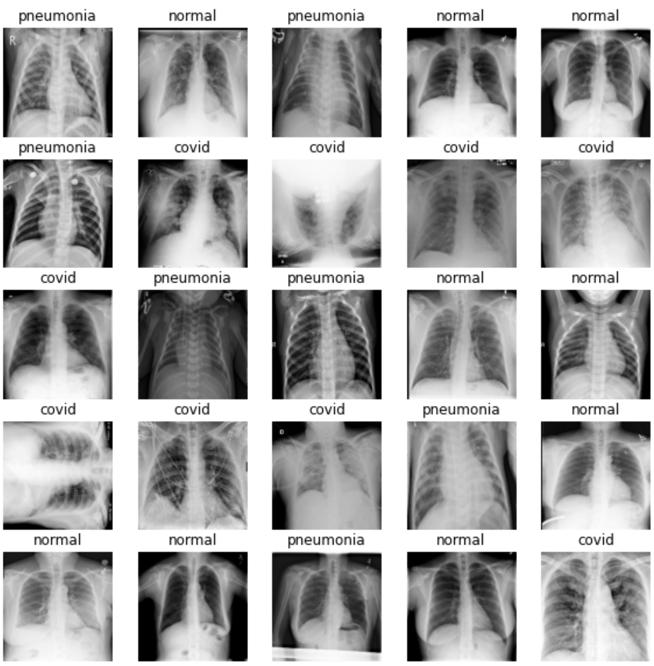


Figure 27: Sample Image Visualization

The data set is divided with testing size of 20%. The split is done with the help of "train_test_split" from the sklearn library.

Table 6: Data Set Split into Training and Testing for Categorical

	Training	**Testing**
Covid-19	545	167
Pneumonia	3875	398
Normal	1341	242

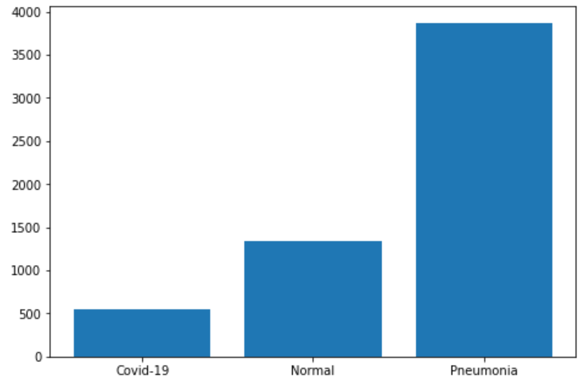


Figure 40: Training Split (Categorical)

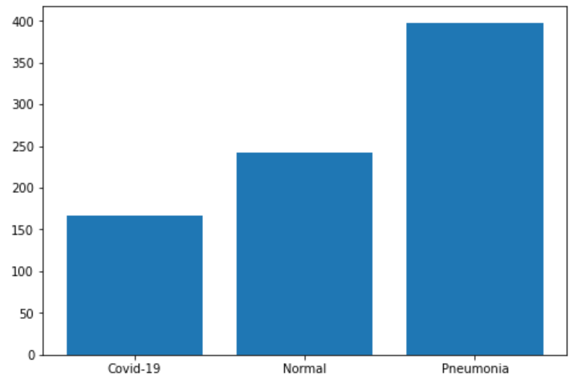
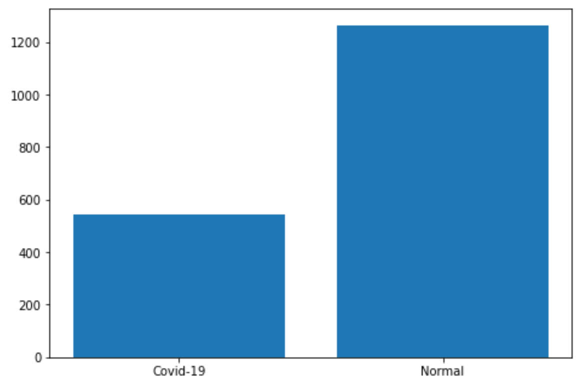


Figure 41: Testing Split (Categorical)

Table 7: Data Set Split into Training and Validation for Binary

	Training	**Testing**
Covid-19	545	167
Normal	1341	242



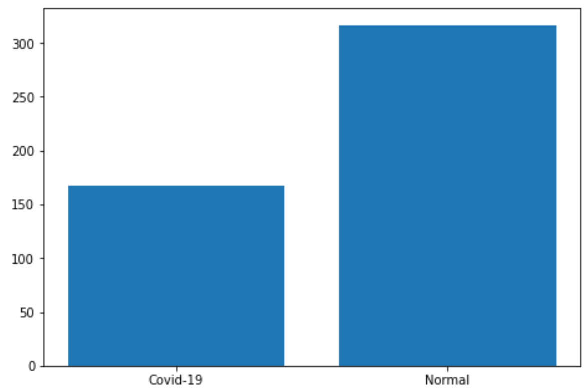


Figure 42: Training Split (Binary) Figure 43: Testing Split (Binary)

The Data is then further saved into a output folder where the dataset is split into the following order.

Table 8: Directory Specification

Categorical Classification Data	**Binary Classification Data**
● Categorical Dataset o train ▪ Normal ▪ Covid ▪ Pneumonia o test ▪ Normal ▪ Covid ▪ Pneumonia	● Binary Dataset o train ▪ Normal ▪ Covid o test ▪ Normal ▪ Covid

These images considered for analysis. Since this is a prediction-based study large sizes in the dataset are required. As this is a medical problem the images are hard to be obtained.

5.3 Image Pre-Processing:

Validation split of 0.1 is performed with a data generator on the training data using the Keras.preprocessing method "ImageDataGenerator" to construct validation and training datasets for model fitting. Along with the data from the tests where augmentation isn't needed. The augmentation aims to minimise data overfitting.

The images we had were scaled down to 256 256 for faster model training, consistent image size, and considering the computational capacity of the device as described in

[6]. From the training data in the dataset folder, the training and validation datasets are generated. With a batch size of 16, the datasets are generated and shuffled. The seed value is 1932, which is chosen at random.

Table 9: Training and Validation Split (Categorical)

	Training	Validation	Testing
[Covid/Normal/Pneumonia]	5186	575	807

Table 10: Training and Validation Split (Binary)

	Training	Validation	Testing
[Covid/Normal]	1449	362	484

Rescaling of 1. /255 is done as referred to [5] and is explained in the below figure.

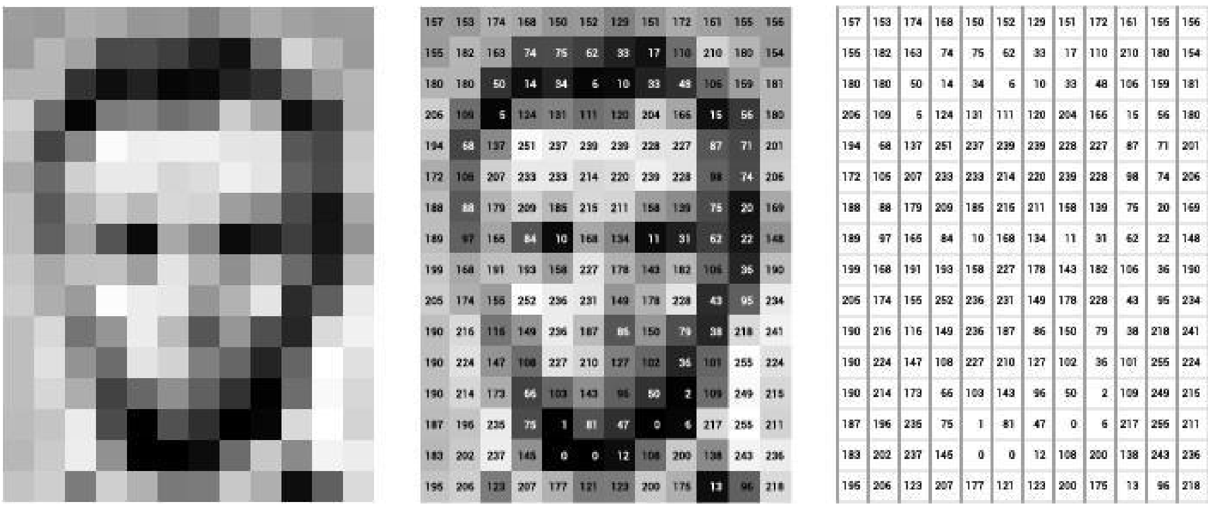


Figure 28: Rescaling Example

With rescaling the values of the pixel which ranges from [0, 255] gets converted into [0,1] allowing us to treat all images in same way. Since X-Rays are based on Grayscale Mechanism the augmentation is made as grayscale colour format reducing the channels of the image.

Now the pre-processed data can be used to fit in a model.

5.3 Image Classification Process:

Image classification process is performed as below in our analysis.

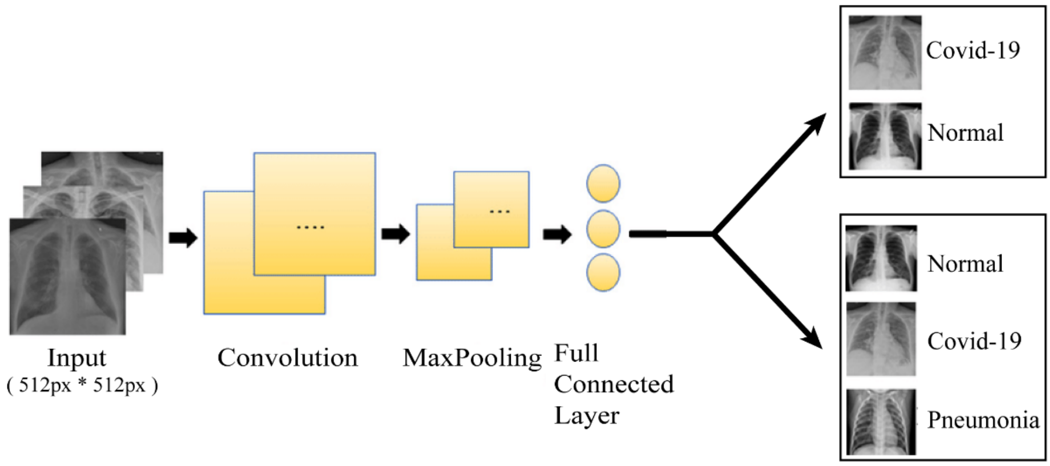


Figure 47: Image Classification

5.5 Experiments:

5.5.1 Preliminary Testing

Several studies are carried out to determine the best CNN efficiency in light of the image database. Covid-19/Normal is used in the experiments. To validate the results using different structures and pre-processing methods, the network architectures with varying numbers of convolutional and completely connected layers, as well as simple image pre-processing techniques were used.

The first structure (CNN#1) was made up of two convolutional layers with 64 and 16 filters, as well as two completely connected (dense) layers with 256 and 1 neurons. In this review, it was the lightest architecture considered.

The second CNN structure is as follows: Three convolutional layers with 256, 128, 64, and 128, 64, 32 filters were used in CNN#2, as well as two entirely connected layers with 128 and 1 neurons. CNN#3 had four convolutional layers (256, 128, 128, and 64 filters) and three completely connected layers, making it the deepest architecture in this review (128, 64, and 1 neuron). For all structures, the filter sizes were set to 45, with a 0.5 dropout for each sheet. Total pooling was used, and 22% pooling was used.

Except for the final convolutional layer of each structure, each layer is considered. **Table 9** show the architectural properties of four considered CNNs.

A total of 5 experiments were performed in this category of COVID-19/Normal, to evaluate and analyse the performance of CNNs under different conditions to achieve an optimal classification of COVID-19 images.

Table 11. Architectural Properties of Four Considered CNNs.

Architecture	CNN Layer	Filters	Filter Size	Pooling and Size	Dropout	Activation
CNN#1	1	128	5*5	Max-Pooling 2*2	0.5	ReLU
	2	32		Max-Pooling 1*1		
CNN#2	1	128	5*5	Max-Pooling 2*2	0.5	ReLU
	2	64				
	3	32				
CNN#3	1	256	5*5	Max-Pooling 2*2	0.5	ReLU
	2	128				
	3	64				
	4	32				
CNN#4	1	32	5*5	Max-Pooling 2*2	0.5	ReLU
	2	64				
CNN#5	1	256	5*5	Max-Pooling 2*2	0.5	ReLU
	2	128				

Based on the above findings, CNN#4 is chosen for preliminary dataset testing to fine-tune the parameters.

For the preliminary tests, a sequential model with two layers depth, filters of (64, 32), filter size of (5,5), and drop out of (0.5) is chosen to avoid overfitting the results. The Adam Optimizer with a learning rate of 0.001 is used since it gives better results. For hidden layers, the activation feature "Relu" is used, with "SoftMax" as the final layer for categorical classification prediction and "Sigmoid" as the final layer for binary classification prediction.

For Adam Optimizer, a simple comparison hyperparameter tuning was performed on different learning rates, with 0.001 being the best learning rate. As a result, 0.001 is chosen as the learning rate.

5.5.2 Test 1: Augmentation vs No-Augmentation:

Most studies in this category preferred to use augmentation to avoid overfitting the results. In contrast, in situations such as x-rays, where detection is based on location, conflicting results could result in False Positives and False Negatives. To validate the scenario, a simplified test is run on the categorical dataset.

5.5.3 Test 2: Image Sizes (256,256) vs (512, 512):

For this analysis, the two image sizes mentioned above are taken into account. The maximum size is (512, 512) because it is the computational limit, while the minimum size is (256, 256) because going below would lose a lot of data from the image because image classification relies heavily on black and white details.

With the results of the above preliminary test, further tests are conducted, and the results are recorded for future reference. We used Transfer Learning Algorithms to improve our prediction performance because we had fewer data and couldn't construct a model from scratch. So, 5 Transfer Learning models are selected based on our study to train which are further used for weighted classifier.

- DenseNet201
- ResNet50
- VGG16
- InceptionV3
- InceptionResNetV2

5.5. Transfer Learning Experiments:

Since the dataset is tiny, Transfer Learning is used to find the model that produces the most accurate results. Since each model has different structure configurations, each model is given a weight to maximise its reliability.

For this experiment, ImageNet weights are used as the model. These models, as state-of-the-art models, provide the baseline accuracy for further classification research.

In a feedforward fashion, DenseNet12113 links each layer to every other layer. The pooling and a fully connected layer follow the initial convolutional layer, and the remaining convolutional layers are followed by the pooling and a fully connected layer. There are 121 layers in all, with over 8 million trainable parameters.

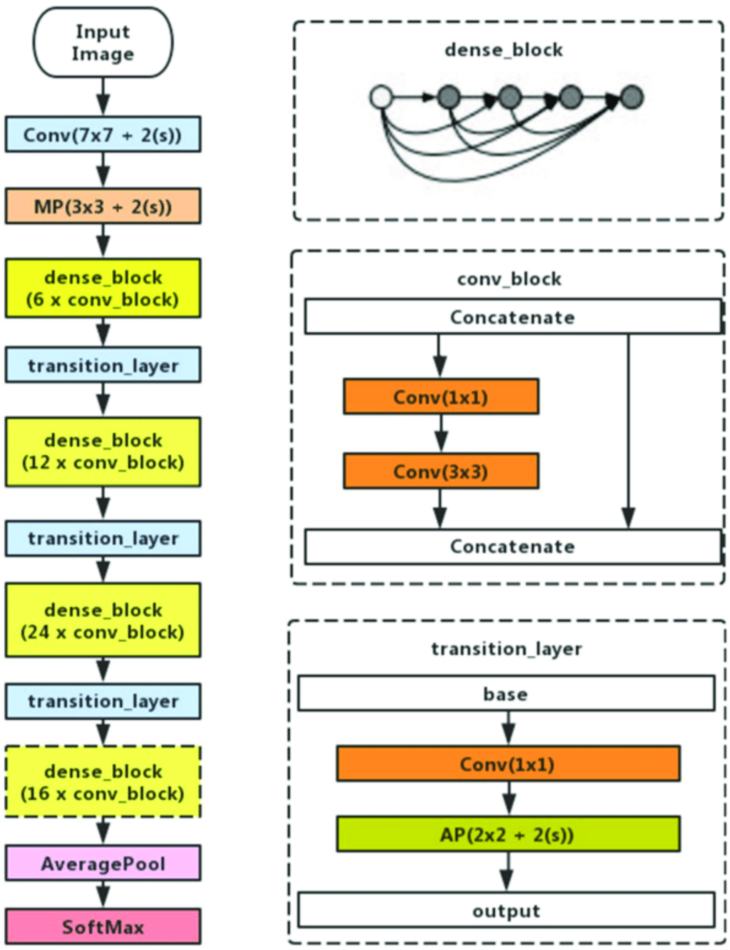


Figure 33: DenseNet121 Structure

ResNet5014 has 50 residual layers, which are designed to address issues including time consumption as the network grows deeper. Its theory is based on identity function skip connections between layers, which improves model accuracy while reducing training time. There are over 23 million trainable parameters in it.

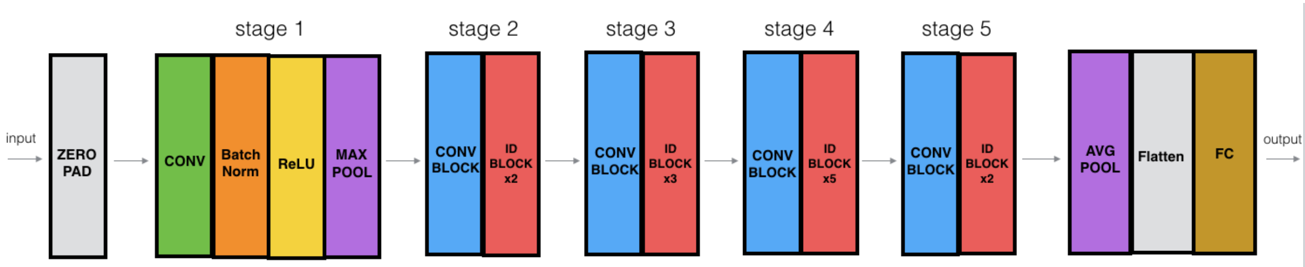


Figure 34: ResNet50 Structure

VGG1615 is a CNN architecture that uses 33 filters and has 16 layers with weights. It has two completely connected layers after the convolutional layers, followed by a SoftMax for performance. For the network, there are approximately 138 million parameters. VGG1913 is similar to VGG16, but it has 19 layers with weights, which gives the network about 143 million parameters.

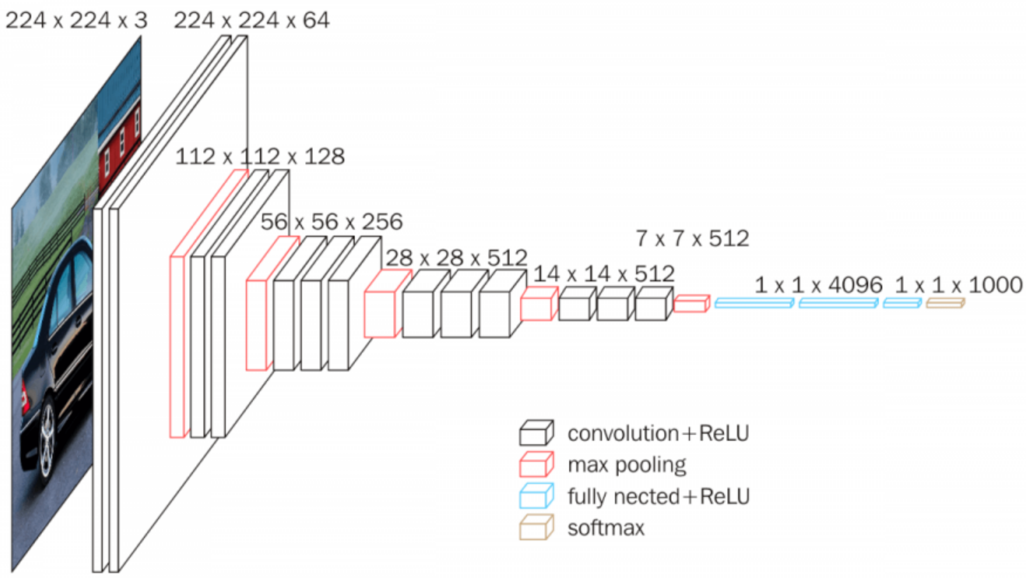


Figure 35: VGG16 Structure

InceptionV3[16] has 42 layers and 24 million parameters. It factorizes convolutions to reduce the number of parameters without decreasing the network efficiency. In addition, novel downsizing was proposed in Inception V3 to reduce the number of features.

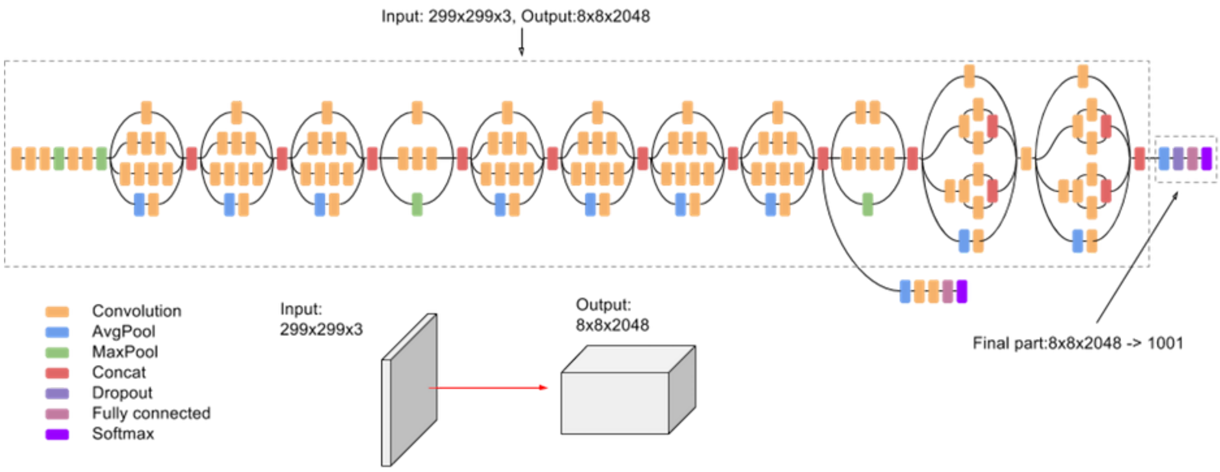


Figure 36: InceptionV3

Inception-ResNet-v2[17] is a convolutional neural network that is trained on more than a million images from the ImageNet database [1]. The network is 164 layers deep and 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 299-by-299.

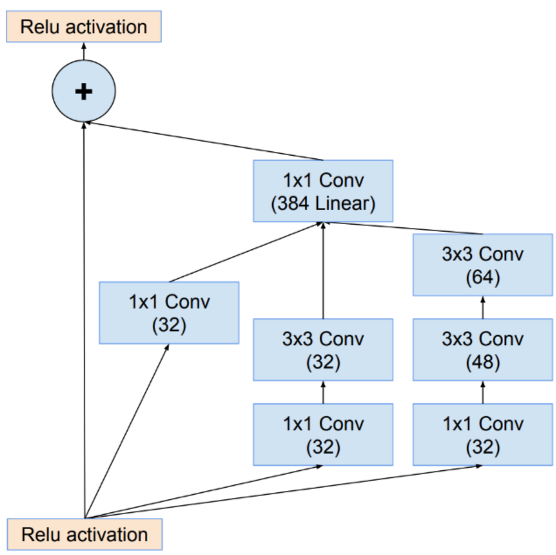


Figure 37: Inception-ResNet-v2 Structure

Each X-ray image was sent to the considered networks with the minimum dimensions required. The performed pre-processing on the considered models to provide consequent images to the models. After training of each model with pre-trained weights, maximum pooling was applied, and features were sent to the fully connected layer (128). The test is performed on each model based the below mentioned scenarios.

5.6 Image Classification:

5.6.1 Scenario-1: Covid19 vs Non-Covid:

In this binary classification the classification is made on Covid vs Non-Covid images on the above transfer learning algorithms.

5.6.2 Scenario-2: Covid19 vs Pneumonia vs Non-Covid

In this categorical classification is made on Covid vs Pneumonia vs Normal are made. Though the bias is present towards the Pneumonia because of the larger dataset of images proper evaluation is done to reduce the false positives and negatives.

5.7 Larger Dataset Experiment:

A simpler test using DenseNet201 is made on a handcrafted larger dataset was used in this experiment.

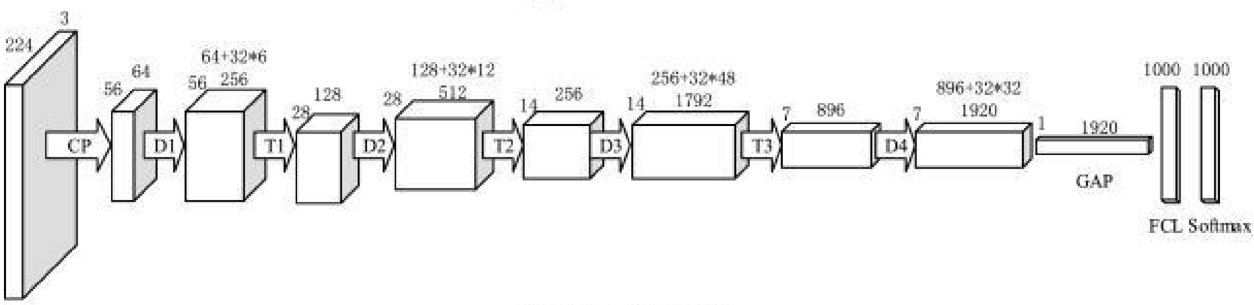
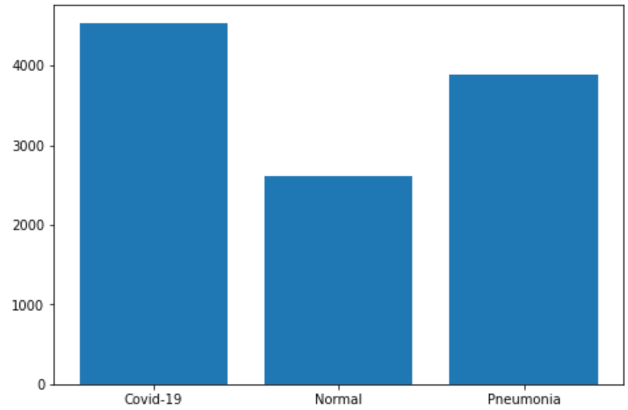


Figure 44: DenseNet201 Architecture

DenseNet-201[18] is a convolutional neural network that is 201 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.



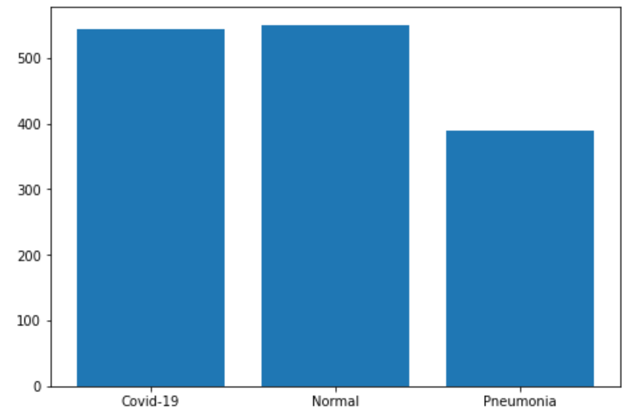


Figure 45: Training Dataset 2

Figure 46: Testing Dataset 2

Testing:	Testing:
TOTAL: 11037	TOTAL: 1486
PNEUMONIA: 3883	PNEUMONIA: 390
COVID19: 4539	COVID19: 545

NORMAL: 2615	NORMAL: 551

This dataset is collective combination of education454 GitHub repository [1.b] and from Kaggle by paultimothymooney which provided pneumonia and normal lung x-rays [1.a] and by Quatar university which provided covid x-rays [1.c]. These 3 datasets are combined manually to create this large dataset and used for prediction analysis.

5.8 Model Evaluation Criteria:

Models can be evaluated using different criteria, such as classification accuracy, Precision, Recall etc. (true positive rate. Accuracy and sensitivity/specificity criterion is not enough, however, especially for imbalanced data; while higher scores can be obtained in one metric, lower scores can be produced by other metrics. By considering all the above-mentioned criteria, AUC was used to evaluate the model performance for the statistical measurement, COVID-19/Normal which had two output classes (labels) and COVID-19/Pneumonia/Normal had 3 output classes(labels). AUC is used to measure the performance of a model. In medical applications, the model with the higher ROC AUC score is more capable of distinguishing between patients with COVID-19 and without COVID-19. "Positive" and "negative" results are the responses of the outputs (classification predictions) obtained from the model. "True" and "false" are the actual data. The accuracy, precision, and recall are calculated as given in

Equation (1), Equation (2), and Equation (3),
respectively:

Accuracy = (TP + TN) / (TP + TN + FP + FN) (1)

Precision = (TP) / (TP+FP) (2)

Recall = (TP) / (TP + FN) (3)

where TP and TN denote the true-positive and true-negative values, respectively; and FP and FN represent false-positive and false-negative values, respectively. 20% and 80% of the data were used for testing and training, respectively. 10% of images are randomly selected for both healthy and coronavirus-infected patients were selected as the validation set.

CHAPTER-6

RESULTS

6.1 CNN Results:

Results of COVID-19/Normal Experiments. In this group, a total of 2295 images (712 COVID-19 and 1583 Normal) were trained in each experiment with the data augmentation procedure, which artificially increases the training samples. The highest accuracy of Experiments 1– 4 was obtained in Exp.4 (98.44%). The highest sensitivity, highest specificity, which is the primary indicator for an imbalanced dataset. Exp.2 and Exp. 3 could not achieve higher rates than Exp.1 and Exp.4 in the evaluation metrics.

Table 12: CNN Experiment Results

Experiment	**Sensitivity [Threshold of 0.75]**	**Accuracy**	**Loss**
CNN#1	100%	97.11%	7.8%
CNN#2	100%	97%	7%
CNN#3	99.07%	95.54%	6.83%
CNN#4	99.39%	99.01%	6.4%
CNN#5	99.79%	97.90	6.13%

The above table shows the obtained results of loss and accuracy along with Sensitivity of the Prediction. Based on the comparison if results Experiment 4 i.e., Convolution Neural Network Model 4 provide a better accuracy and results with minimal loss value.

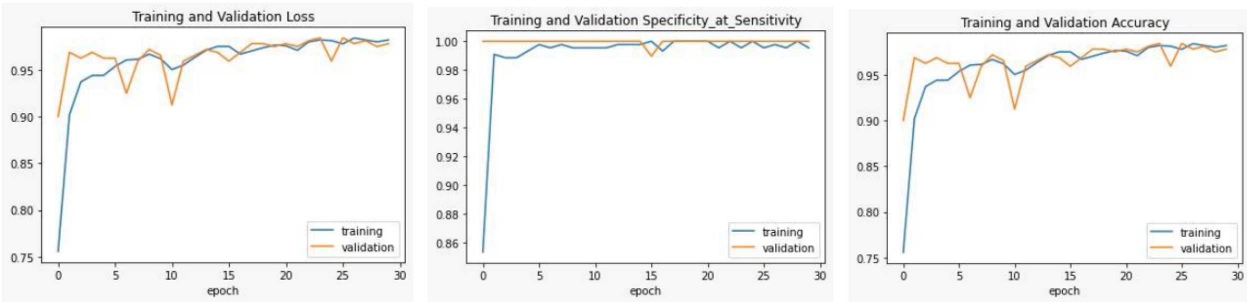


Figure 38: CNN#4 Result Graphs

6.2 Results of Test 1:

Table 13: Augmentation Prediction

	Accuracy	Precision	Recall	Epochs
	Training/ Testing	Training/ Testing	Training/ Testing	
Augmented	0.88/0.76	0.90/0.77	0.89/0.75	11
Not Augmented	0.96/0.80	0.96/0.80	0.95/0.78	11

The Non-Augmented providing better results compared to Augmented results in terms of accuracy, precision and recall values.

6.3 Results of Test 2:

Table 14: Image Size Prediction

Since the Non-Augmented data provided better results, we chose Non-Augmented data and variable image sized to choose which provides us the better results.

Image Size	Accuracy	Precision	Recall	Epochs
	Training/ Testing	Training/ Testing	Training/ Testing	
(256, 256)	0.95/0.79	0.96/0.77	0.95/0.75	5
(512, 512)	0.87/0.78	0.89/0.81	0.84/0.88	5

Despite getting better accuracy for smaller image, it has better precision and recall value compared to smaller image size.

6.4 Results of Transfer Learning Experiments:

Accuracy, Precision, Recall and AUC (Area under curve) are considered as metrics for the evaluation of each model. Finally, F1-Score is calculated to find a better model for prediction.

6.4.1 Scenario 1: Covid-19 vs Normal

Results of COVID-19/Normal Experiments. In this group, a total of 2295 images (712 COVID-19 and 1583 Normal) were trained in each experiment without the data augmentation procedure. The images were then split into training and testing data.

Table 15: Transfer Learning Results – Scenario 1

Table 15.1 Training:

Model Name	Accuracy	Loss	Precision	Recall	F1-Score	AUC
DenseNet121	99.11	0.07	0.9908	99.68	0.995	98.85
ResNet50	69.89	0.68	0.69	1	0.79	50
VGG16	99.9	0.003	1	0.99	0.99	99
InceptionV3	99.9	0.05	0.99	0.99	0.992	99
Inception-ResNet-V2	99.7	0.11	0.99	0.99	0.996	99

Table 15.2 Testing:

Model Name	Accuracy	Loss	Precision	Recall	F1-Score	AUC
DenseNet121	98.3	0.67	0.975	1	98.75	97.9
ResNet50	68.26	0.68	0.654	1	0.79	50
VGG16	98.76	0.02	98.44	0.99	0.996	99.9
InceptionV3	98.97	1.72	0.99	0.99	99.2	98.9
Inception-ResNet-V2	99.59	0.10	0.993	1	0.996	99.2

Though we were able to obtain high accuracy results there is a significant amount of loss present in the testing data evaluation and also there is a significant amount of validation loss obtained through the model training. This is due to the overfitting of data in the process of training the model. Out of the obtained results VGG16 seems to provide better predictions with lower amount of loss for the binary classification.

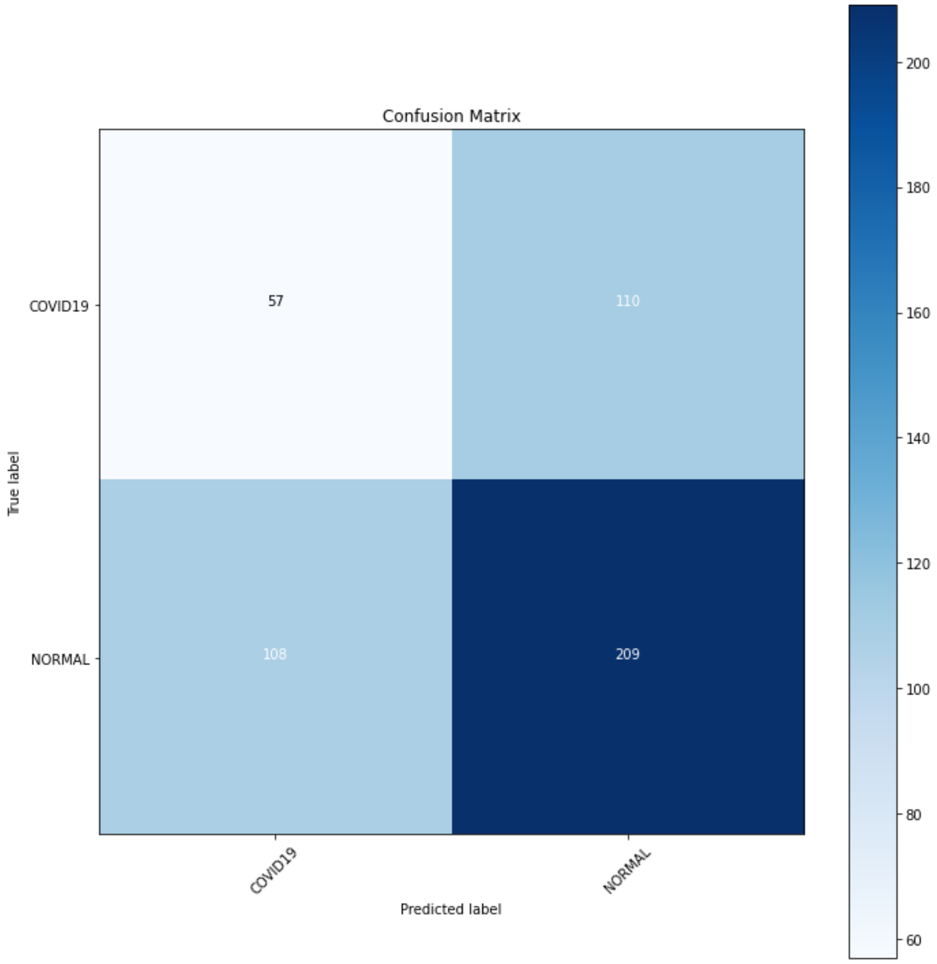


Figure 38: Confusion Matrix – VGG16

6.4.2 Scenario 2: Covid-19 vs Pneumonia vs Normal

Results of COVID-19/Pneumonia/Normal Experiments. In this group, a total of 5568 images (712 COVID-19, 4273 Pneumonia and 1583 Normal) were trained in each experiment without the data augmentation procedure. The images are further split into training and testing data.

Table 16: Transfer Learning Results – Scenario 2

Table 16.1 Training:

Model Name	Accuracy	Loss	Precision	Recall	F1-Score	AUC
DenseNet121	98.9	0.24	0.99	0.98	0.985	98.94
ResNet50	77	0.9	0.776	0.769	0.71	90
VGG16	98.8	0.24	0.98	0.988	0.98	99.3
InceptionV3	96.59	0.19	0.965	0.965	0.789	98
Inception-ResNet-V2	99.51	0.02	0.995	0.995	0.995	99.9

Table 16.2 Testing:

Model Name	Accuracy	Loss	Precision	Recall	F1-Score	AUC
DenseNet121	79	0.62	0.81	0.78	0.82	92
ResNet50	71.62	1.12	0.718	0.712	0.71	87.93
VGG16	86.25	8.9	0.86	0.863	0.862	90.06
InceptionV3	78.9	4.04	0.79	0.789	0.789	85.75
Inception-ResNet-V2	82.28	3.2	0.823	0.82	0.823	88.98

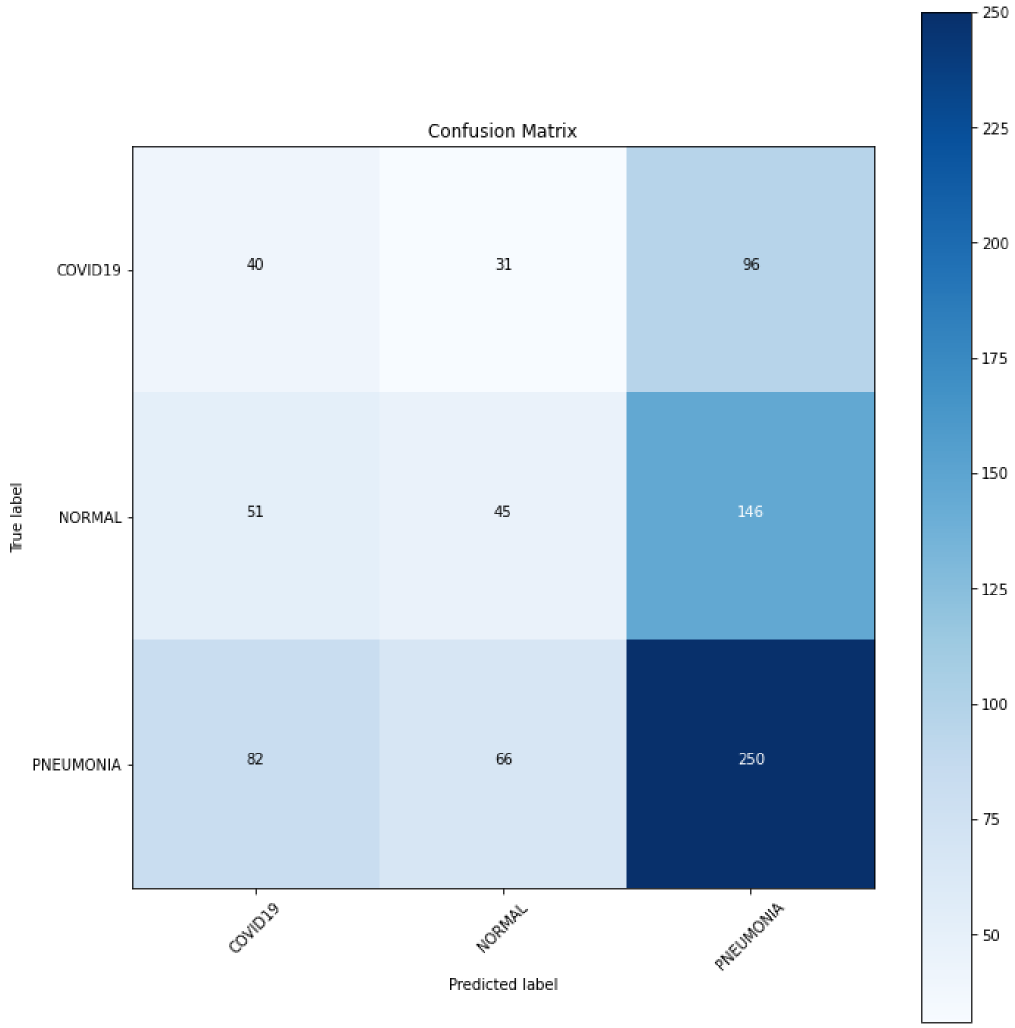


Figure 39: Confusion Matrix – DenseNet121

Despite obtaining high accuracy results there is a significant amount of loss present in the training and testing data and there is a significant amount of validation loss obtained through the model training. This is due to the overfitting of data in the process of training the model. Out of the obtained results DenseNet121 seems to provide better predictions with lower amount of loss for the categorical classification.

There are issues like overfitting of data increasing the loss of the model. Further tests should be made on the model by freezing and unfreezing required layers in Transfer Learning methods. The lower image quality and inconsistent images resulted in increase of False positives and Negatives. Manual of checking of proper images and collection of such data will help in increasing the quality analysis of the research. Since the study depends on the minimal chances of detecting covid by variating from the pneumonia the bias was obvious causing the categorial and binary classification biased towards pneumonia or normal classes. To reduce the bias in the model we need to increase the Covid-19 data, augmenting the data could reduce the data quality and reduce the minimal chance of detecting covid-19.

Implementing methods like Weighted Classifier and Ensemble Learning could be used to obtain better classification of the model. Larger dataset with classified images and of better quality will be needed to reduce false positives and negatives for the results.

6.5 Results of Larger Dataset Experiment

With the handcrafted dataset we tested a Transfer Learning Algorithm such as DenseNet201 for the following test where we have done about 5 epochs and obtained the following results.

Table 17: Training and Testing larger dataset

DenseNet201	**Accuracy**	**Loss**	**Precision**	**Recall**	**F1-Score**	**AUC**
Training	87.6	0.3	0.889	0.86	0.874	99.05
Testing	83.6	0.38	0.84	0.83	0.81	95.6

Thought the accuracy is low, the prediction showed better results than that of the previous tests made.

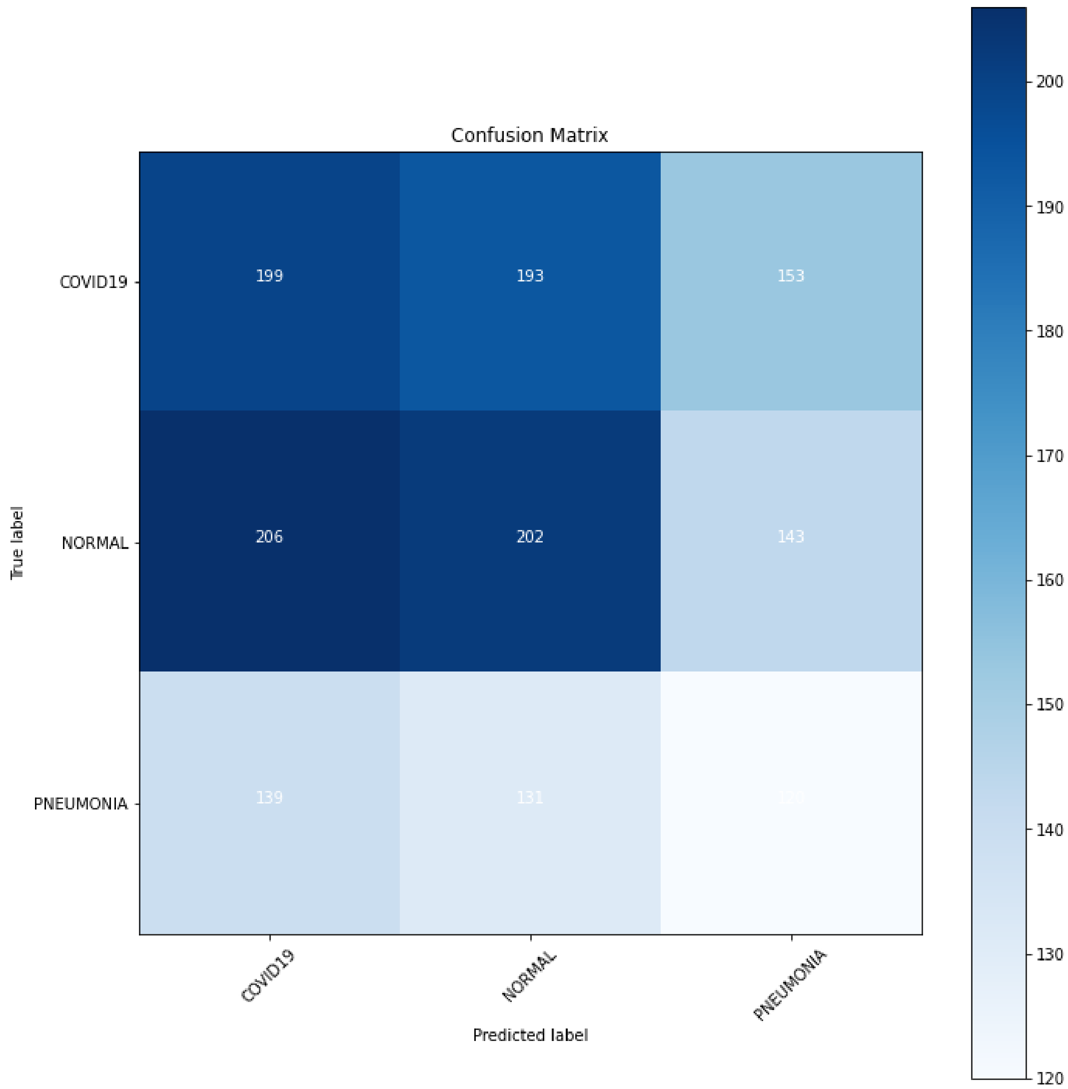


Figure 47: Confusion Matrix – Categorical – Larger DS

Though there are fair share of false positives and false negatives, it also can be noticed that the true positives and negatives increased in this scenario where larger datasets is used. This further proves the point of having inconsistent data is not reliable. The above result can be improved further by fine-tuning the model and improving the quality of the database. As the images in the dataset have both Posterior-Anterior (PA) and Anterior-Posterior (AP) where we only need Posterior-Anterior for our classification.

CHAPTER-7

THE USER INTERFACE MODEL

7.1 Header

Visual design is similar to user interface design when you are viewing or conveying information visually. As a result, a user interface (UI) is needed to make the model more useful and interactive. As a result, a UI built with the Flask framework and a Flask backend makes the user and the model we build much more interactive.

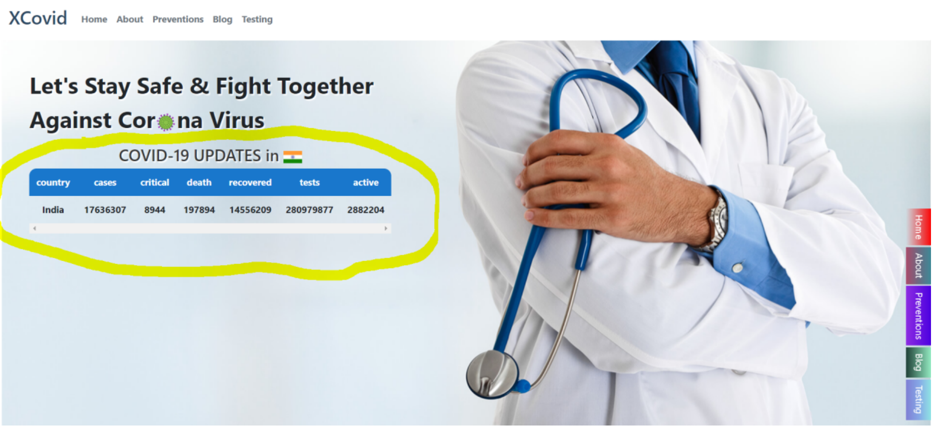


Figure x: Header

This is the home page, and the data presented in the table changes with the update with the change in confirmation of data.

7.2 Fetch Code:

```
fetch(`https://corona.lmao.ninja/v2/countries/india`)
.then((response)=>{
    return response.json();
})
.then((data)=>{
    document.getElementById("flag").src = data.countryInfo.flag;
    document.getElementById("country").innerHTML = data.country;
    document.getElementById("cases").innerHTML = data.cases;
    document.getElementById("critical").innerHTML = data.critical;
    document.getElementById("death").innerHTML = data.deaths;
    document.getElementById("recoverd").innerHTML = data.recovered;
    document.getElementById("tests").innerHTML = data.tests;
    document.getElementById("active").innerHTML = data.active;
})
```

Figure x: Fetching Cases

The data comes from an API, from which we fetch the country flag, country name, cases in that country, critical patients count, number of deaths, number of recovered, number of active cases etc.

7.3 Information

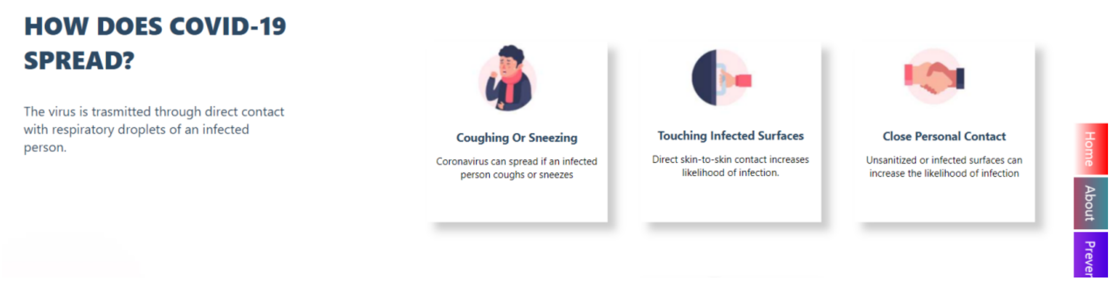


Figure x: Helpers

This page is about how does that covid-19 spread as the virus is transmitted through direct contact with respiratory droplets of an infected person.

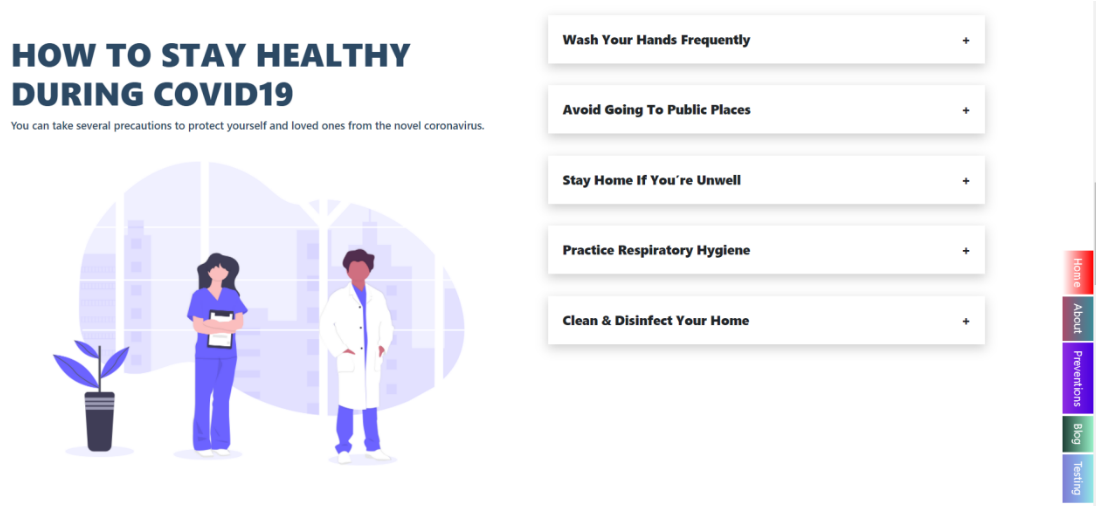


Figure x: Suggestions

This component is about prevention or measurements to be taken to avoid spread of corona. The prevention tips are given through the box, each box has a message inside it which will be displayed once we press the plus button.

Figure x: Working Suggestions Sample

An accordion is used to show (and hide) HTML content.

7.4 Social Media Updates

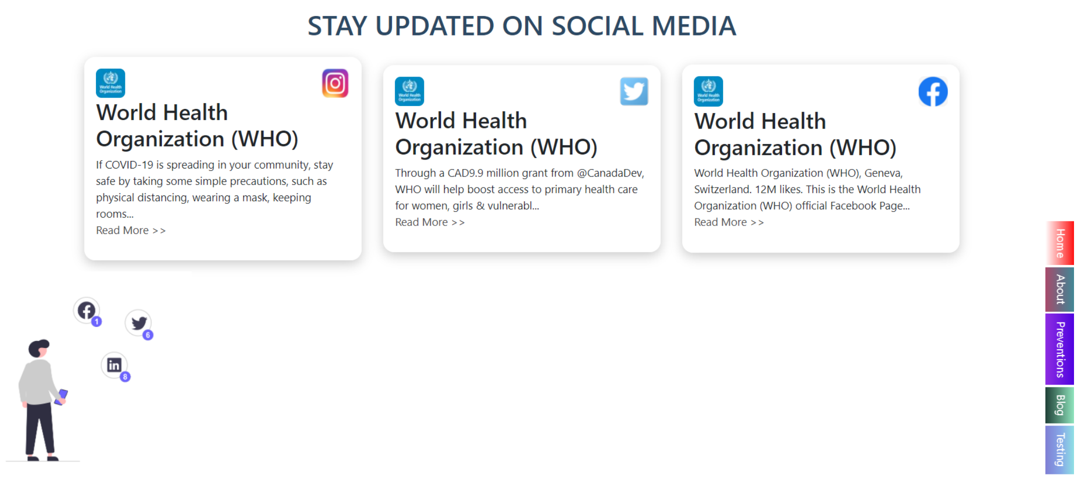


Figure x: Social Media Helpers

Social media is everywhere, we can figure out the news from social media updates during this pandemic time. This component keeps us updated about the prevention and measurements that are to be taken in this pandemic time

7.5 Prediction Analysis

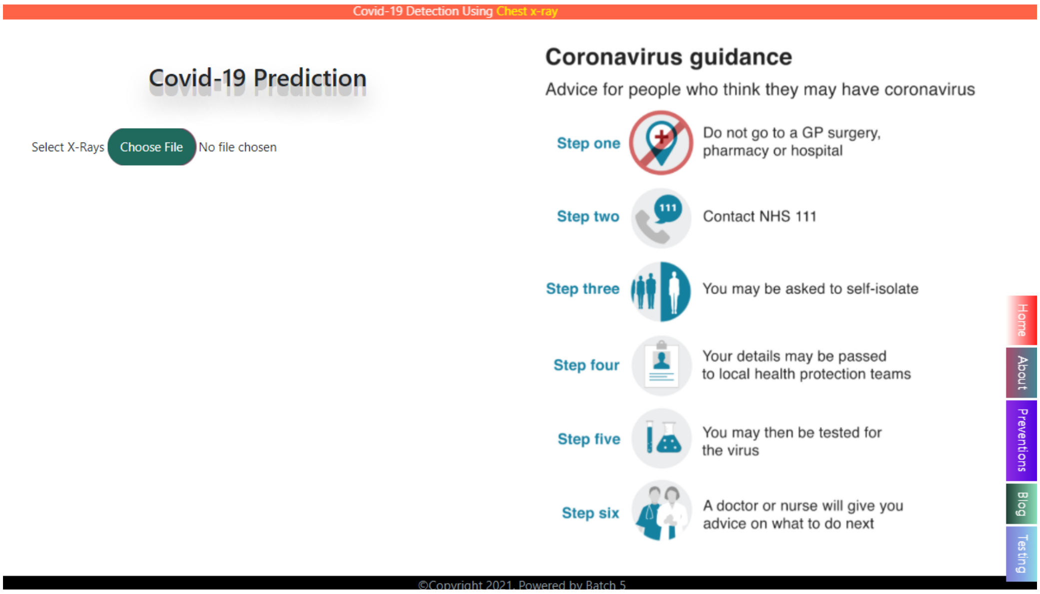


Figure x: Upload X-Rays

Finally, the prediction part is done here and we will be choosing an x-ray image

The main goal in our User interface is that to make the website Responsive so that our website can be more interactive with the users.

So, the view is made responsive to be made open in different formats.

7.6 View in tablet

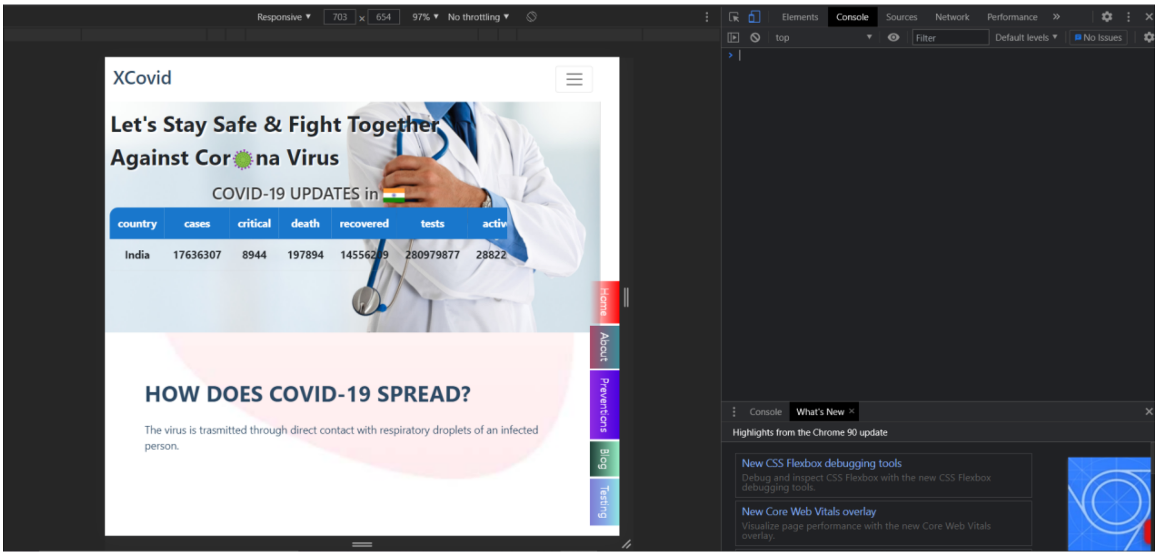


Figure x: Tablet View

7.7 View in Mobile

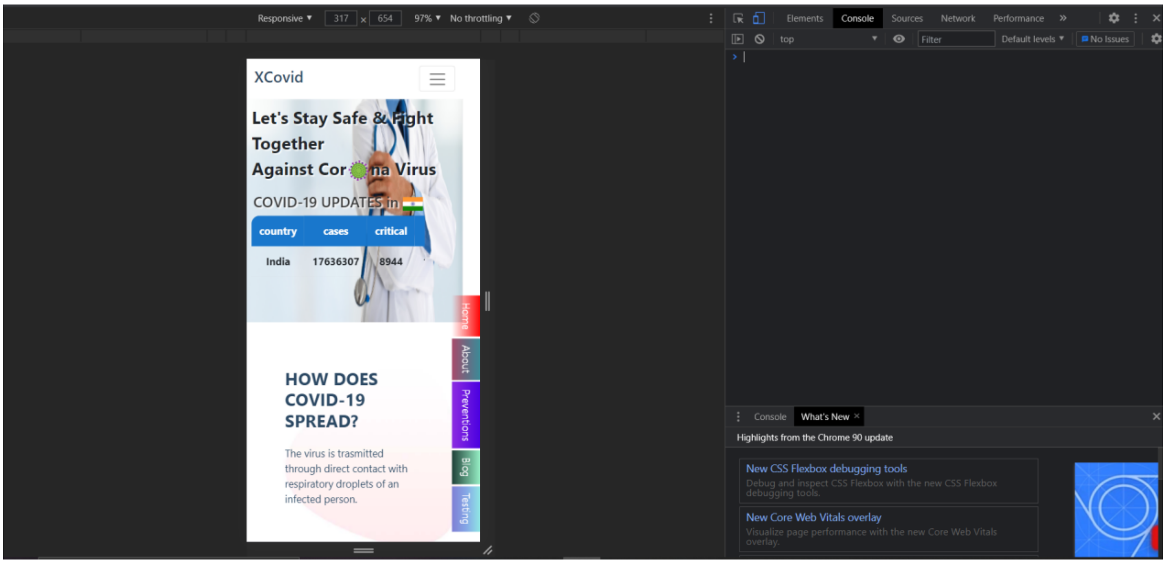


Figure x: Mobile View

CHAPTER-8

CONCLUSION

The COVID-19 pandemic is now wreaking havoc on the global population's health. The most important step in prevention or eradication is to foresee it, followed by treatment. Since the lung scans, we are using here may have fibrosis, which can be solved by image classification, prediction is a major task in this case.

We've implemented some deep learning algorithms to predict the covid-19 in this example. COVID-19 detection from chest X-ray images is important for both doctors and patients in terms of reducing testing time and costs. Artificial intelligence and deep learning may recognise images for the tasks that have been taught. Several experiments were carried out in this study to detect COVID-19 in chest X-ray images with high precision using CNNs. COVID-19/Normal and COVID-19/Pneumonia/Normal were considered for the classification. Different image dimensions, different network architectures, state-of-the-art pre-trained networks, and machine learning models were implemented and evaluated using images.

Since there is fewer data available, the forecast could be accurate, but it is not reliable enough to be used in a clinical trial. More data and computational power are needed to produce more accurate and useful results. Though our current project might not be completely trustworthy, further research may show that it is more useful and trustworthy. This research would take more time and effort to complete. Overfitting and lower precision values are found in the sample, which should be subjected to quality analysis to produce more reliable results. The inconsistent databases resulted in false predictions as well as smaller working time resulted in hurried results.

With more data available, a better analysis of image classification and Covid-19 prediction may be possible. This also helps us in getting early care for patients with pneumonia.

The findings of this study can be used as a starting point for further research.

CHAPTER-9

FUTURE SCOPE

The model's results were skewed due to the scarcity of data. Instead of settling with fewer data, further analysis can be done with a greater amount of data using other techniques such as oversampling or SMOTE methods. A simplified model could also be created with the increased computational capability to achieve better performance. With further testing, a better model with a larger dataset can be created from scratch and made clinically accessible.

To design a better model, a larger dataset of higher-quality images must be found. The accuracy of the model can be improved with further changes to the dataset. Fine Tuning the transfer learning models will also create better prediction models. With smaller computational ability we had to compromise over the smaller image size and data.

Also, it appeared that the Larger Dataset provided better performance, which can be further improved by fine-tuning the models and improving the consistency of the dataset and algorithm pre-processing.

Ensemble learning can be used to improve accuracy and decrease the number of false positives and negatives.

CHAPTER-10

REFERENCES – BIBLIOGRAPHY

1. 1.1 Chest X-ray images (Pneumonia) Dataset, Accessed at: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

   1.2 Education454 Dataset accessed at https://github.com/education454/datasets
2. COVID-19: The first documented corona virus pandemic in history by Yen-Chin Liu,Rei-Linkuo,Shin-Ru-Shihat https://www.sciencedirect.com/science/article/pii/S2319417020300445
3. Effects of social distancing on the spreading of Covid-19 inferred from mobile phone data by Hamid Khataee, Istvan Scheuring, Andras Czirok and Zoltan Neufeld at https://www.nature.com/articles/s41598-021-81308-2
4. Personal Protective Equipment: Challenges and Strategies to Combat COVID-19 in India: A Narrative Review by Neeraj Sharma, Zubeda Hasan & Annop Velayudhanat https://journals.sagepub.com/doi/full/10.1177/0972063420935540
5. Real-Time Reverse Transcription- Polymerase Chain Reaction Assay for SARS-associated Coronavirus at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3322901/
6. Identifying COVID-19 from Chest CT images: A Deep Convolutional Neural NetworksBasedApproach at https://www.hindawi.com/journals/jhe/2020/8843664/
7. Pneumonia at https://www.mayoclinic.org/diseases-conditions/pneumonia/symptoms-causes/syc-20354204

8. Novel Feature Selection and Voting Classifier Algorithms for COVID-19 Classification in CT Images by IEEE

9. Deep Learning-based detection and analysis od COVID-19 on chest X-ray images

10. Based Identifying COVID-19 from Chest CT Images- A Deep Convolutional Neural Networks Approach

11. XCOVNet: Chest X-ray Image Classification for COVID-19 Early Detection Using Convolutional Neural Networks.

12. ImageNet. http://www.image-net.org

13. https://www.researchgate.net/figure/Left-DenseNet121-architecture-Right-Dense-block-conv-block-and-transition-layer_fig4_331364877

14. https://www.sciencedirect.com/science/article/pii/S2210670720308179

15. https://neurohive.io/en/popular-networks/vgg16/

16. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/44903.pdf

17. https://paperswithcode.com/method/inception-resnet-v2

18. https://www.mathworks.com/help/deeplearning/ref/densenet201.html