# Pneumonia Detection via Chest X-Rays


## Group Members

Muhammad Basit Ali(01-131172-017)

Uzair Abdullah Mir (01-131172-047)

***Supervisor:*** Engr. Joddat Fatima



A Final Year Project submitted to the Department of Software Engineering, Faculty of Engineering Sciences, Bahria University, Islamabad in the partial fulfillment for the award of degree in Bachelor of Software Engineering

**July 2021**

# THESIS COMPLETION CERTIFICATE

Student Name: **Uzair Abdullah Mir** Enrolment No: **01-131172-047**

Student Name: **Muhammad Basit Ali** Enrolment No: **01-131172-017**

Programme of Study: Bachelor of Software Engineering

Project Title: **Pneumonia Detection via Chest X-Rays**

It is to certify that the above students' project has been completed to my satisfaction and to my belief, its standard is appropriate for submission for evaluation. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index at 14% that is within the permissible limit set by the HEC. I have also found the thesis in a format recognized by the department.

Supervisor's Signature: ___________________________

Date: 6/9/2021 Name: Joddat Fatima

# Certificate of Originality

This is certify that the intellectual contents of the project "**Pneumonia Detection via Chest X-Rays"** are the product of my/our own work except, as cited properly and accurately in the acknowledgements and references, the material taken from such sources as research journals, books, internet, etc. solely to support, elaborate, compare, extend and/or implement the earlier work. Further, this work has not been submitted by me/us previously for any degree, nor it shall be submitted by me/us in the future for obtaining any degree from this University, or any other university or institution. The incorrectness of this information, if proved at any stage, shall authorities the University to cancel my/our degree.

Name of the Student: **Uzair Abdullah Mir**

Signature: ________________________ Date: 08-06-2021

Name of the Student: **Muhammad Basit Ali**

Signature: ________________________ Date: 08-06-2021

# Abstract

Although pneumonia is a treatable disease, it kills more children than any other disease or infection. According to a UNICEF report, Pneumonia claimed lives of 808,694 children under the age of five in 2017[1]. In Pakistan, an estimate of 58,000 children died of Pneumonia in 2018 placing Pakistan at 3rd number in most deaths by pneumonia in the world [1]. The major cause of deaths by pneumonia in Pakistan is low availability of doctors and medical personals in certain cities and areas. So, we are developing a web-based application that can detect pneumonia from the image of a patient's chest (lungs) x-ray. With unavailability of doctors in certain regions, patients need to travel long distances to get doctor appointments wasting crucial time. This project is aimed to assist patients to get diagnosis fast from home and act relatively. We are using deep learning algorithm called Convolutional Neural Network (CNN) that will be trained and tested on a dataset of chest x-rays. Users would simply upload their chest x-ray image on the website and the CNN algorithm will predict if the patient has pneumonia or not from the x-ray image. With even low availability of doctors in this covid pandemic this project can really help patients to get diagnosed in time as patient's condition keeps getting critical with time. Not all cases of pneumonia are serious, in some cases the patient has mild symptoms so in that case they should not panic but in case of moderate and severe pneumonia, the patients need to seek immediate medical assistance so that treatment can be started in time. Therefore, the deep learning model will then predict its severity as "Mild", "Moderate" or "Severe". The website will use the users' current location from their GPS and use that location to find pulmonologists near their current location and display their details with location of the pulmonologists using Google Maps API so that the user may make appointments for further diagnosis and treatment. Let us hope the project helps patients as we think it would.

**Keywords:** Pneumonia, Deep Learning, Convolutional Neural Networks (CNN)

# Dedication

*We would like to dedicate this project to our beautiful and beloved parents who not only have supported us financially bus also blessed us with their moral support.*

# Acknowledgments

*Our deepest gratitude to Allah Almighty who has granted us courage, patience, knowledge, and skills that were necessary to carry out this project and our degree.*

*We would like to thank our supervisor, Ma'am Joddat Fatima for guiding us through our journey of this project and giving her expert opinion and pointing out our mistakes and helping us to complete this project. Her contribution and constructive criticism pushed us to make our work as good as we could.*

*Our utmost regards to our parents who have made it all possible after Allah Almighty, without their support we would never have achieved anything that we have today.*

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Our project solves a real life problem by helping people getting diagnosed from the comfort of their home by detecting presence of pneumonia from their chest x-ray images.

## 1.1. Motivation

Pakistan ranked 3rd in deaths caused by Pneumonia in 2018 with 58,000 deaths in the world [1]. Our project aims at assisting patients that cannot find medical personnel in time with a system that can predict the presence of pneumonia with high precision and save crucial time wasted in travelling to find medical personnel just for diagnosis.

Moreover, the recent outbreak of Coronavirus has taken the whole world by surprise. It has taken more than 3,500,000 lives worldwide. The chances of getting infected by covid are highest in hospitals. This project will also prevent the unnecessary time spent in hospitals to get diagnosed, reducing the chance of catching covid-19 virus.

## 1.2. Problem statement or research questions

Pneumonia is a treatable disease but still it kills a lot of people worldwide most of which are children. *"Every day, nearly 2,200 children under the age of five die from pneumonia, a curable and mostly preventable disease."* said Henrietta Fore, Executive Director of UNICEF [1]. One of the main reasons for these deaths is late detection of pneumonia in patients. Pneumonia keeps getting critical if not treated in time and eventually result in casualties. With unavailability of doctors in certain regions, patients need to travel long distances to find doctor appointments which is waste of crucial time. So, this project is aimed to assist patients to get diagnosis fast and act relatively.

## 1.3. Objectives

The main objectives that we have focused on achieving are:

1. Make people able to get diagnosed from anywhere through internet.
2. Reduce the need to visit hospitals to reduce crowd in hospitals in this pandemic.

## 1.4. Main contributions

**What is new, different, better, significant?**

Although there has been some research done on pneumonia detection using AI but there is no web or mobile application developed yet that would help ordinary patients and doctors to diagnose pneumonia from chest x-rays.

**Why is the world a better place because of what I have done?**

Recent breakout of covid pandemic has caused many causalities around the globe due to its property of fast and easy spread. The chances of getting infected by covid are highest in hospitals as there are patients present day and night. This app limits the need to visit hospitals, minimizing the chances of getting infected by covid or any other contagious disease.

**What is now known/possible/better because of this project?**

This project makes it possible to get diagnosed from your nearest internet service point if you have digital image of your chest x-ray which is nowadays easily available.

## 1.5. Report organisation

The structure of the report is as follows.

- **Chapter 1:** In this chapter, we discussed our motivation for this project and the problem statement that we were working on. Furthermore, we discussed the objectives of the project and how is this project going to help people.
- **Chapter 2:** In this chapter we discuss the background and history of our project. Past work done in our related field is discussed and examined.
- **Chapter 3:** In this chapter we discuss system requirements. It includes gathering requirements and organizing them by allocating them into functional and nonfunctional requirements.
- **Chapter 4:** In this chapter we have included the graphical representation of our project that will allow users and developers to get better picture of what is being implemented and how will it be implemented.
- **Chapter 5:** This chapter includes implementation of the system. It contains different working strategies and methodologies to develop the system.
- **Chapter 6:** This chapter gives the details of testing including unit, integration and system testing to make sure that application is working fine.
- **Chapter 7:** This chapter concludes the project and report and talks about how this project can be further improved and enhanced.

# Chapter 2

# Background Study/Literature Review

Advancements in the performance of deep learning algorithms and datasets has enabled Artificial Intelligence to overtake the performance of medical professionals in different areas of medical imaging tasks like diabetic retinopathy detection [2], skin cancer classification [3], arrhythmia detection [4] and haemorrhage identification [5].

Specifically, CNN's application has been studied on pulmonary diseases over the years and new research is being done by the top organizations of the world. Automated (or unsupervised) diagnosis of pulmonary diseases and chest/lungs related conditions using deep learning algorithms has been given extra attention recently. Pulmonary tuberculosis classification [6] and lung nodule detection [7] are two of many research paper published in the recent years.

## 2.1. Related Work

| Published by | Year Published | Description |
|---|---|---|
| Khalid El Asnaoui ,Youness Chawki, and Ali Idri. | 2020 | This team of researchers did an extensive research on "Automate Detection and Classification of Pneumonia based on Chest X-Rays Images using deep Learning techniques". Their research also included detection of other pneumonia like diseases like SARS, COVID-19 and MERS [8]. |
| Pranav Rajpurkar and Jeremy Irvin. | 2017 | This group of researchers made a deep neural network and named it CheXNet. CheXNet [9] is a 121 layered CNN model which was developed by a group of researchers in 2017. It was used to detect multiple pulmonary disease including pneumonia. They achieved an accuracy of 0.7680 for pneumonia. |

Although this is one on the best research paper on pulmonary disease detection but like every piece of work done in this topic, it lacks one thing that we offer, an actual (web) application that allows radiologists and ordinary patients to detect pneumonia from their digital chest x-ray.

# Chapter 3

# System Requirements

### 3.1. Use Case Diagram

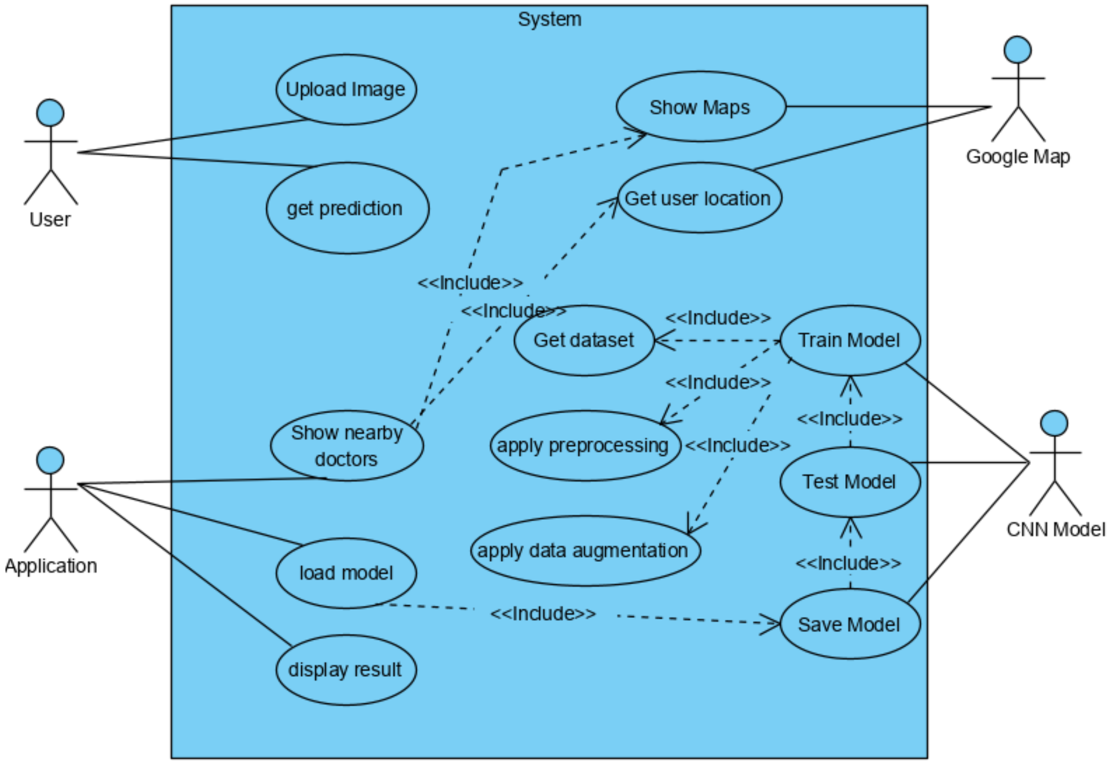


*Figure 1: System Use case Diagram*

#### 3.1.1. Use case Description:

**Upload Image:**

The user will use upload Image button to upload their chest x-ray image from their device. Once uploaded the image will be displayed on the screen.

**Get Prediction:**

The user will get the result by clicking on get Prediction button.

**Train Model:**

CNN model is trained on the dataset using transfer learning technique. (This use case is performed only once)

**Save Model:**

The weights of trained model are saved for later use. (This use case is performed once)

**Load Model:**

The saved model is loaded on the web to predict user uploaded image(s).

**Get User Location:**

Google Maps API will access user’s current location.

### Show Maps:

Google Maps API is used to display google maps on the website.

### Show nearby Doctors:

Using the user’s current location, pulmonologists are show on the Google maps near him in case the user want to make an appointment for further treatment/diagnosis.

### Display Result:

The website will show either positive or negative for pneumonia.

## 3.2. Functional Requirements

**Upload Image:**

<table>
<tr><td colspan="2">Use Case ID:</td><td colspan="2">UC-1</td></tr>
<tr><td colspan="2">Use Case Name:</td><td colspan="2">Upload Image</td></tr>
<tr><td colspan="2">Actor(s):</td><td colspan="2">User</td></tr>
<tr><td colspan="2">Pre-Conditions:</td><td colspan="2">User must have Image on PC and connected to Wi-Fi</td></tr>
<tr><td colspan="2">Priority:</td><td colspan="2">High</td></tr>
<tr><td colspan="2">Basic Flow:</td><td colspan="2">User uploads Image by clicking on “Upload Image” button</td></tr>
<tr><td colspan="3">Actor Actions</td><td>System Response</td></tr>
<tr><td>1</td><td colspan="2">User clicks on “Upload Image” button and selects the x-ray image from the directory on computer</td><td>2 System checks if uploaded image is in the correct format.</td></tr>
<tr><td colspan="3">Alternative Course of Action (if any)</td><td></td></tr>
<tr><td colspan="3">Actor Action</td><td>System Response</td></tr>
<tr><td></td><td colspan="2"></td><td>2.a if the image is not in the correct format, the system itself converts the image in the correct format.</td></tr>
</table>

*Table 1: Functional Requirement 1*

**Get Prediction:**

<table>
<tr><td colspan="2">Use Case ID:</td><td colspan="2">UC-2</td></tr>
<tr><td colspan="2">Use Case Name:</td><td colspan="2">Get Prediction</td></tr>
<tr><td colspan="2">Actor(s):</td><td colspan="2">User</td></tr>
<tr><td colspan="2">Pre-Conditions:</td><td colspan="2">User must have uploaded chest x-ray Image.</td></tr>
<tr><td colspan="2">Priority:</td><td colspan="2">High</td></tr>
<tr><td colspan="2">Basic Flow:</td><td colspan="2">User clicks on “Predict” button and the System predicts pneumonia from the uploaded image</td></tr>
<tr><td colspan="3">Actor Actions</td><td>System Response</td></tr>
<tr><td>1</td><td colspan="2">User clicks on “Upload Image” button and selects the x-ray image from the directory on computer</td><td>2 System checks if uploaded image is in the correct format.</td></tr>
<tr><td>3</td><td colspan="2">User clicks on “Predict” button</td><td>4 system predicts pneumonia from the uploaded image</td></tr>
</table>

*Table 2: Functional Requirement 2*

**Display Result**

| | | |
|---|---|---|
| **Use Case ID:** | UC-3 | |
| **Use Case Name:** | Display result | |
| **Actor(s):** | Application | |
| **Pre-Conditions:** | User must have uploaded chest x-ray Image and clicked on “Predict” button. | |
| **Priority:** | High | |
| **Basic Flow:** | User uploads Image by clicking on “Upload Image” button and the System returns the result if pneumonia is detected or not and displays the prediction result | |
| **Actor Actions** | | **System Response** |
| **1** | User clicks on “Upload Image” button and selects the x-ray image from the directory on computer | **2** System checks if uploaded image is in the correct format. |
| **3** | User clicks on “Predict” button | **4** system predicts pneumonia from the uploaded image |
| | | **5** system displays the result of prediction |

*Table 3: Functional Requirement 3*

**Show Maps**

| | | |
|---|---|---|
| **Use Case ID:** | UC-4 | |
| **Use Case Name:** | Show Maps | |
| **Actor(s):** | Google Maps | |
| **Pre-Conditions:** | Google Maps API is connected with the application | |
| **Priority:** | Medium | |
| **Basic Flow:** | User uploads Image by clicking on “Upload Image” button and the System returns the result if pneumonia is detected or not. If detected, google maps will be accessed | |
| **Actor Actions** | | **System Response** |
| **1** | User clicks on “Upload Image” button and selects the x-ray image from the directory on computer | **2** System checks if uploaded image is in the correct format. |
| **3** | User clicks on “Predict” button | **4** system predicts pneumonia from the uploaded image |
| | | **5** if pneumonia is detected, application will access google maps from google maps API |
| **Alternative Course of Action (if any)** | | |
| **Actor Action** | | **System Response** |
| | | **5.a** if pneumonia is not detected, application will not access google maps. |

*Table 4: Functional Requirement 4*

**Get User Location**

| | | | |
|---|---|---|---|
| **Use Case ID:** | | UC-5 | |
| **Use Case Name:** | | Get user location | |
| **Actor(s):** | | Google Maps | |
| **Pre-Conditions:** | | User is connected to Wi-Fi and has activated his/her location | |
| **Priority:** | | Medium | |
| **Basic Flow:** | | User uploads Image by clicking on "Upload Image" button and the System returns the result if pneumonia is detected or not. If detected, application will get user's current location | |
| **Actor Actions** | | | **System Response** |
| **1** | User clicks on "Upload Image" button and selects the x-ray image from the directory on computer | | **2** System checks if uploaded image is in the correct format. |
| **3** | User clicks on "Predict" button | | **4** system predicts pneumonia from the uploaded image and displays the result if detected or not |
| | | | **5** if pneumonia is detected, google maps API will access user's current location |
| **Alternative Course of Action (if any)** | | | |
| **Actor Action** | | | **System Response** |
| | | | **5.a** if pneumonia is not detected, google maps API will not access user's current location. |

*Table 5: Functional Requirement 5*

**Show nearby Doctors**

| | | | |
|---|---|---|---|
| **Use Case ID:** | | UC-6 | |
| **Use Case Name:** | | Show nearby Doctors | |
| **Actor(s):** | | Application, User | |
| **Pre-Conditions:** | | Pneumonia is detected from the image uploaded by the user | |
| **Priority:** | | Medium | |
| **Basic Flow:** | | User uploads Image by clicking on "Upload Image" button and the System returns the result if pneumonia is detected or not. If detected, details of nearby pulmonologists will be shown with respect to user's current location | |
| **Actor Actions** | | | **System Response** |
| **1** | User clicks on "Upload Image" button and selects the x-ray image from the directory on computer | | **2** System checks if uploaded image is in the correct format. |
| **3** | User clicks on "Predict" button | | **4** system predicts pneumonia from the uploaded image and displays the result if detected or not |
| | | | **5** if pneumonia is detected, application will show location of nearby pulmonologists with respect to user's current location |
| **Alternative Course of Action (if any)** | | | |
| **Actor Action** | | | **System Response** |
| | | | **5.a** if pneumonia is not detected, application will not show any doctors' details |

*Table 6: Functional Requirement 6*

### 3.3. Interface Requirements

#### 3.3.1. User Interface:

The user will interact with our website by first uploading his/her chest x-ray image which will be displayed on the screen. User can also re-upload the image. Once the correct image is uploaded, the user can get the result/prediction for the chest x-ray.

#### 3.3.2. System Interface:

For the frontend, we have used simple HTML, CSS and JavaScript and for the backend, we have used .Net Core and C#.

### 3.4. Database Requirements

Since the system does not ask users for registration, it means that it does not store any users' data. It only stores X-Ray files for a small amount of time and then junks it because our system does not need to store it. So the database of our system is not extensive.

The database used is SQL server database and it can be upgraded to Azure database if the requirements change in the future and the project is deployed on a big scale.

### 3.5. Non-Functional Requirements

#### 3.5.1. Security

The system provides complete security to the user data. The personal data of the patient is completely secure the website does not collect any data. It only takes the X Ray File and user location. The user location is not stored, it is just taken and then junked after the process.

#### 3.5.2. Performance

The system has efficient performance and does not lag or slow down at any point. There are no performance issues.

#### 3.5.3. Usability

The system is very easy to use and the user experience is designed keeping view the ease of usability.

### 3.5.4. Modifiability

There is always room for modification in the project and our project is completely modifiable.

### 3.5.5. Reliability

The system is very reliable and has no availability issues. It has efficient code and the server is optimized for minimal downtime.

### 3.5.6. Technological Constraint:

The deep leaning algorithm cannot give a 100% accuracy, which means that there is a chance for the prediction to be wrong but it is very minor.

## 3.6. Project Feasibility

### 3.6.1. Technical Feasibility:

This feasibility analysis helped us determine whether the technical resources meet capacity and whether the technical team is capable of converting the ideas into working systems.

We calculated all the costs and technical requirements for this project keeping in view the market trends and demands.

Our system is not very complicated and vast and it requires minimal operational costs and technical resources. Our team is capable to scale this project on a bigger scale with little to technical obstacles.

### 3.6.2. Operational Feasibility:

This feasibility analysis helped us determine if our project is good enough to be operational on a corporate level. We analysed if our project can be set operational on a bigger scale and serve clients.

We reached out to various medical and corporate firms and discussed our project with them and after an extensive discussion we were came to a point that our project has good operational scope but to be operational on industrial level we need to work more on the accuracy of the prediction model and increase the datasets.

### 3.6.3. Legal Feasibility

This feasibility analysis helped us determine if our project is compatible in the market keeping in view the legal policies and regulations.

We did a quick overview of the medical legal policies and came to a point that our project is compliant with all the legal policies because we never recommend any medications or treatment. Our system just predicts the pneumonia and then refers the patients to pulmonologists nearby.

### 3.6.4. Ethical Feasibility

This feasibility analysis helped us determine if project is compatible keeping in view the ethical concerns of the society.

We talked to various people around us regarding our project and also some doctors and ensured that our project does not violate any ethical value.

## 3.7. Analysis Models

### 3.7.1. Activity Diagram

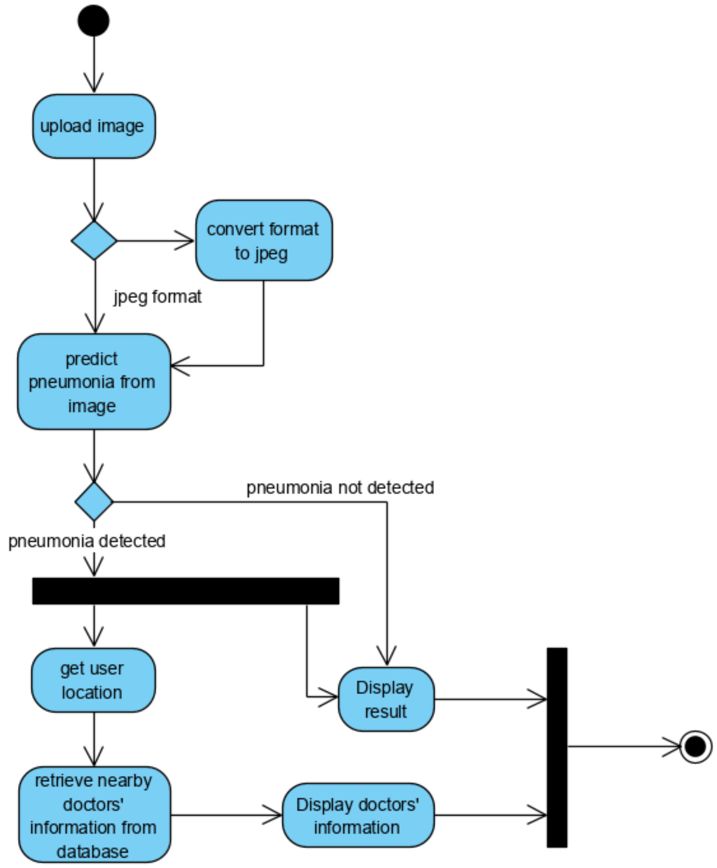


*Figure 2: Activity Diagram*

Our project solves a real-life problem and brings ease into the life of pneumonia patients to some extent. Although it can’t be perfect and 100% accurate, but still it fulfils the requirements of the market and people.

This project has very simple use cases and has a very user-friendly experience. All the functional and non-functional requirements are fulfilled and the project is feasible for the market.

# Chapter 4

# System Design

### 4.1.1. Design Approach

We are using a Function Oriented Design Approach for our project. Our project is focused of the function of Pneumonia Detection, which is further divided into sub functions like “uploading image”, “Applying pre-processing”, “Training model” and “Testing Models” etc. The main function is divided in many sub functions which all perform certain tasks to perform the main task. A top-down approach is used where the main feature/function is divided into sub functions.

## 4.2. Design Constraints [Optional]

Biggest Constraint is high cost and high time consumptions as each module/requirement needs to be planned, designed and tested separately. Function oriented design can get messy as it does not have a proper structure like Object Oriented Design.

## 4.3. System Architecture

For this system, we used MVC architecture as it is one of most efficient architectures for state-of-the-art applications. Our system has three modules and it fits perfectly in the MVC architecture. We put the database in MODEL, all the functions including image uploading, pneumonia prediction in CONTROLLER and all the web pages in VIEW.

## 4.4. Logical Design

### 4.4.1. Class Diagram

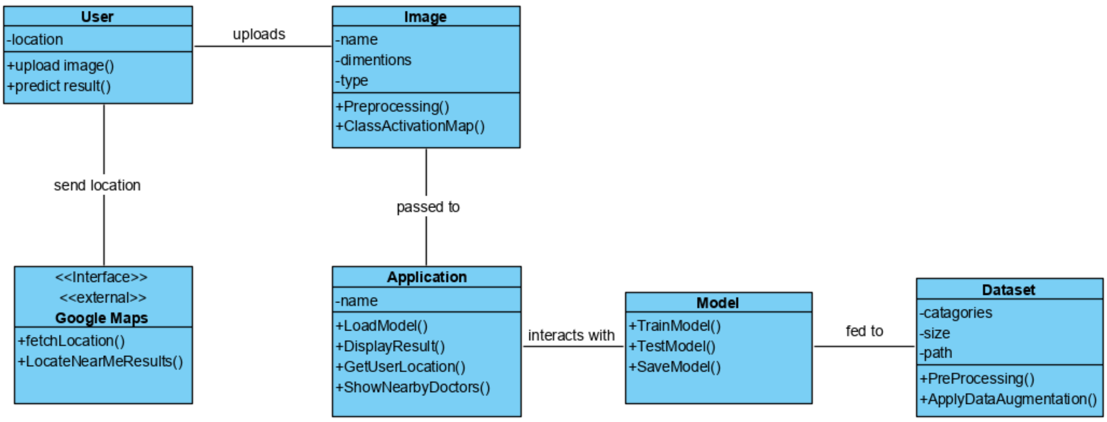

Figure 3: System Class Diagram

### 4.4.2. Data Flow Diagram (DFD)

#### *4.4.2.1. DFD Level 0*

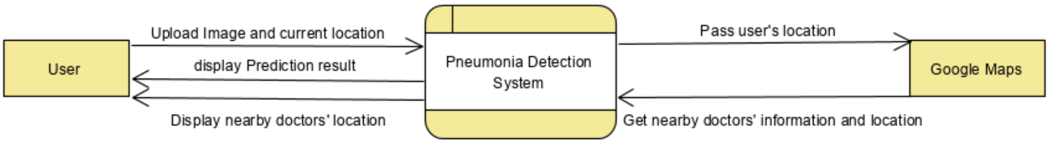


*Figure 4: DFD Level 0*

#### *4.4.2.2. DFD Level 1*

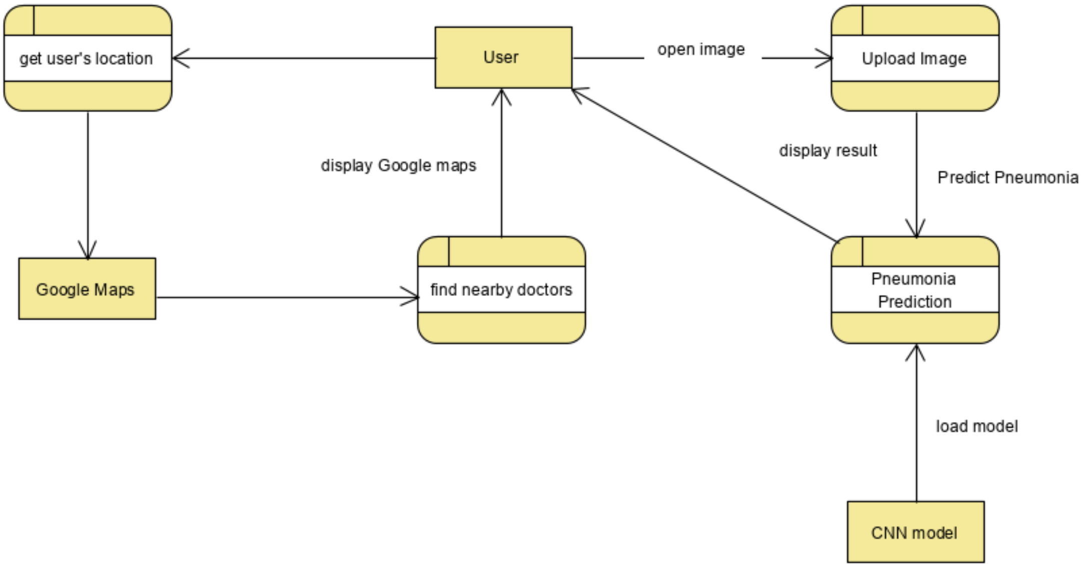


*Figure 5: DFD Level 1*

### *4.4.2.3. DFD Level 2*

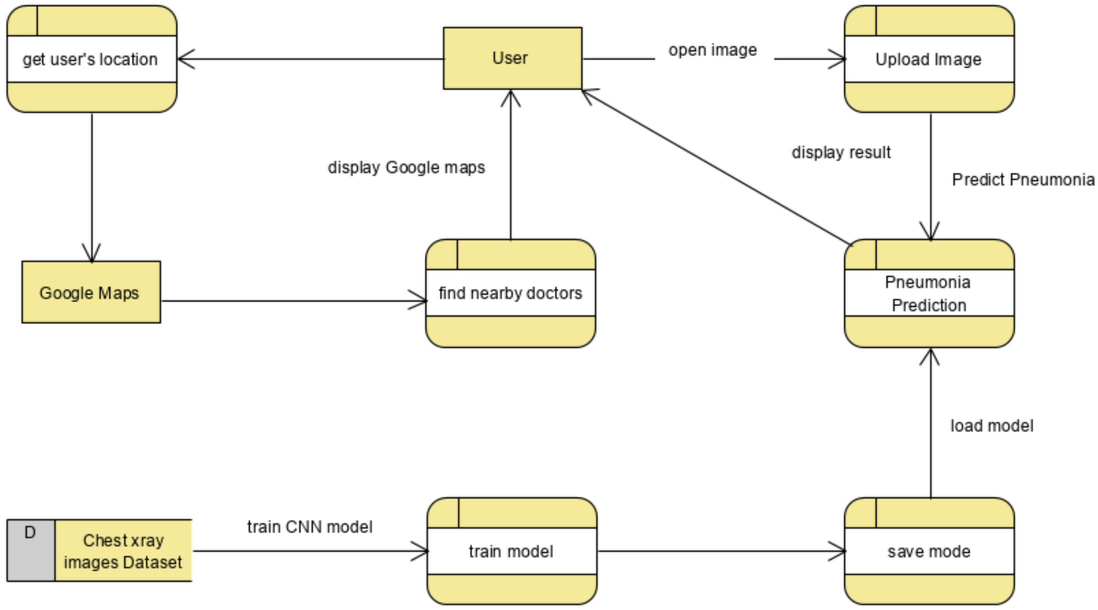


*Figure 6: DFD Level 2*

## 4.5. Dynamic View

### 4.5.1. Sequence Diagram

### *4.5.1.1. Model Preparation*

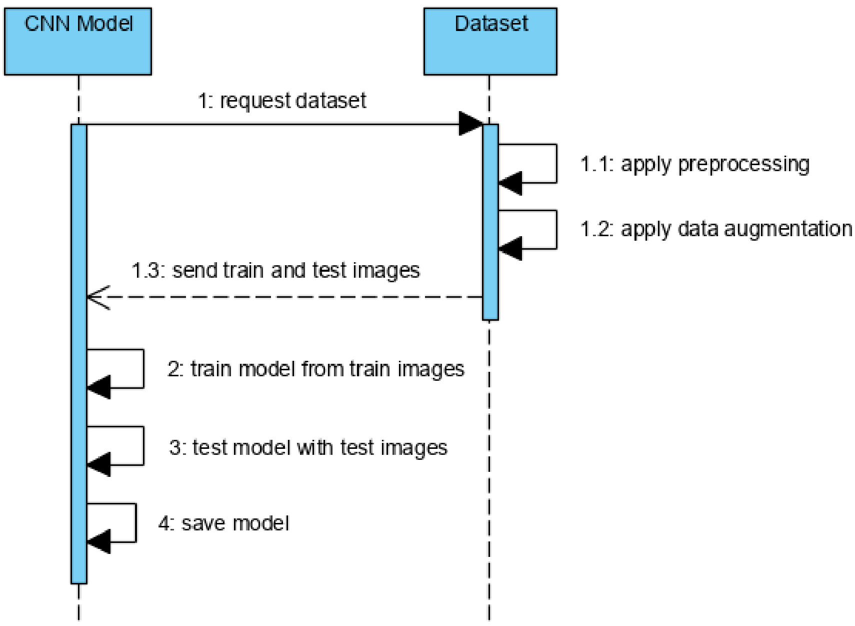


*Figure 7: Model Preparation Sequence Diagram*

### *4.5.1.2. Pneumonia Detection*

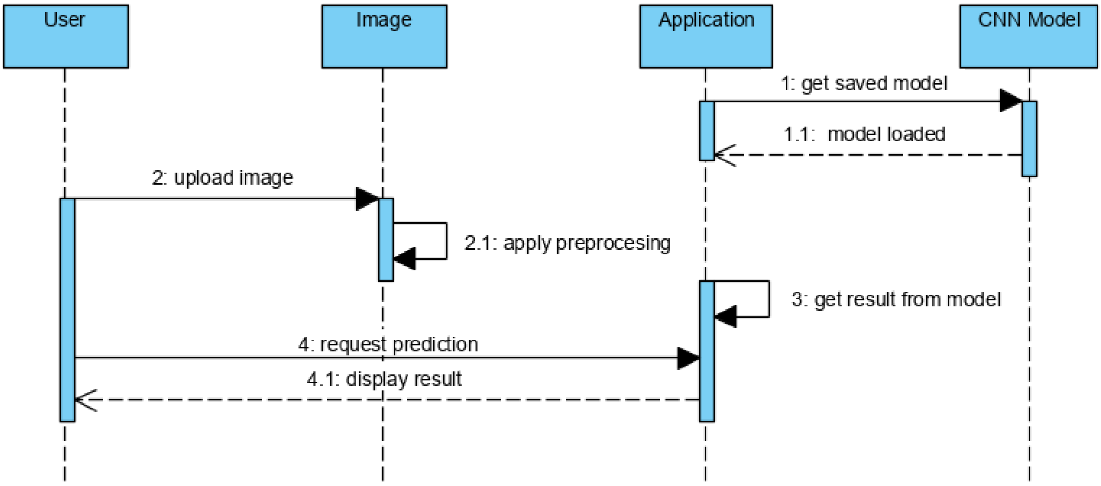


*Figure 8: Pneumonia Detection Sequence Diagram*

### *4.5.1.3. Google Maps display*

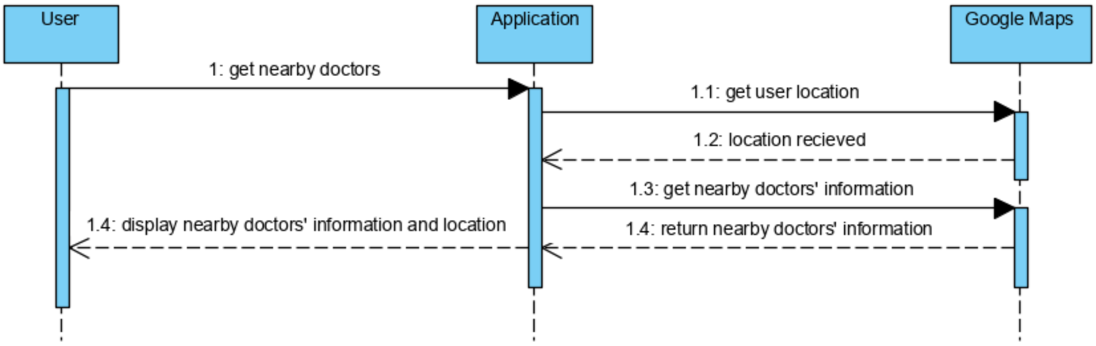


*Figure 9: Google Maps Display Sequence Diagram*

### *4.5.1.4. Overall Sequence Diagram*

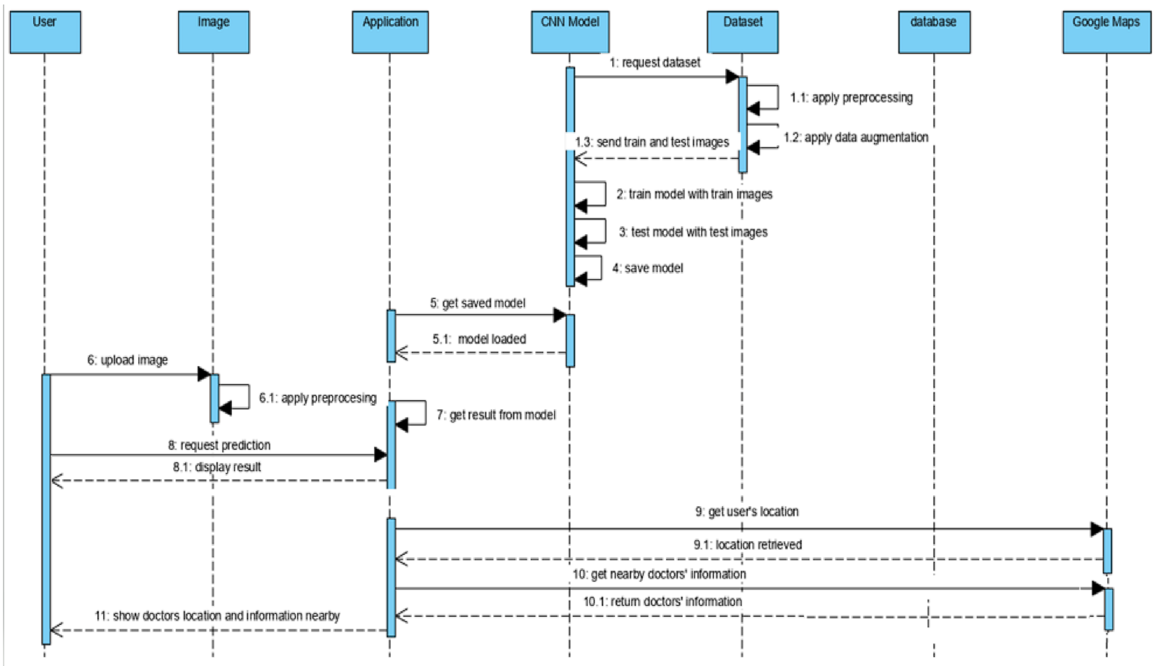


*Figure 10: Full project Sequence Diagram*

## 4.6. Component Design

### 4.6.1. Component Diagram

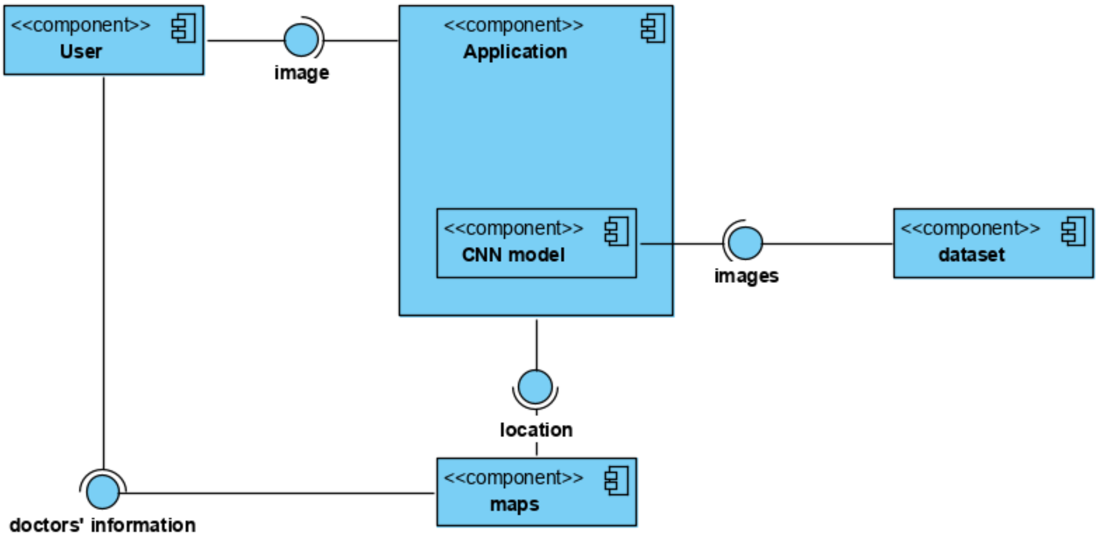


*Figure 11: Component Diagram*

### 4.6.2. Deployment Diagram

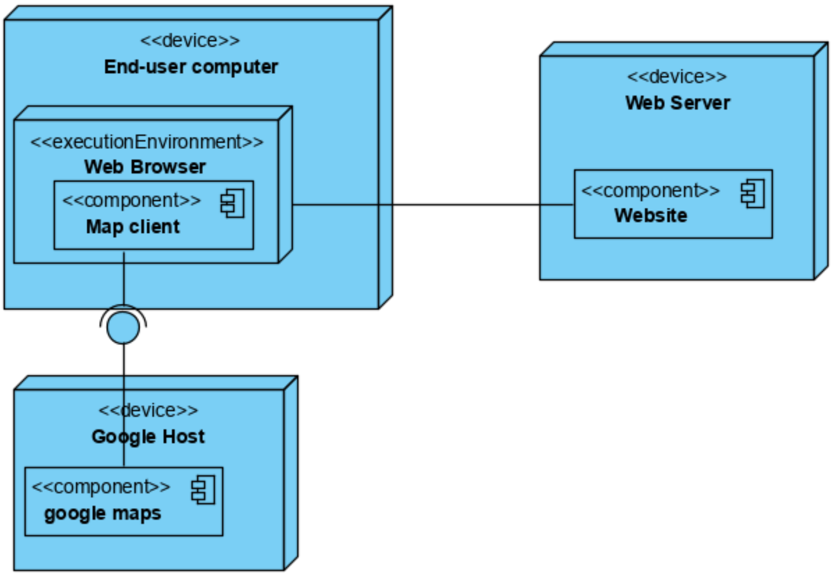


*Figure 12: Deployment Diagram*

## 4.7. Data Models

### 4.7.1. Entity Relation Diagram (ERD)

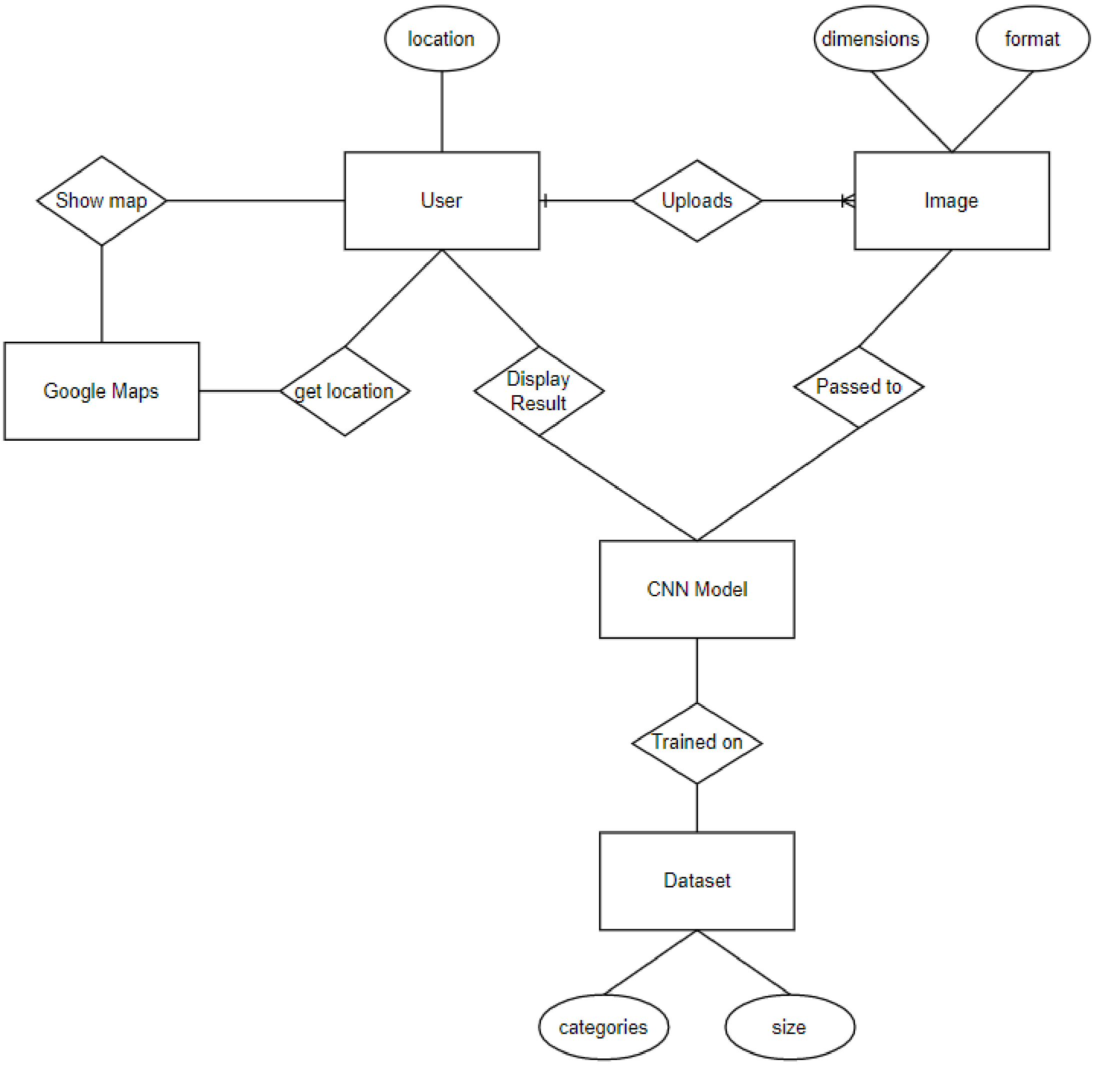


*Figure 13: ER Diagram*

## 4.8. User Interface Design

### 4.8.1. Low fidelity Prototype

**Home Page:**

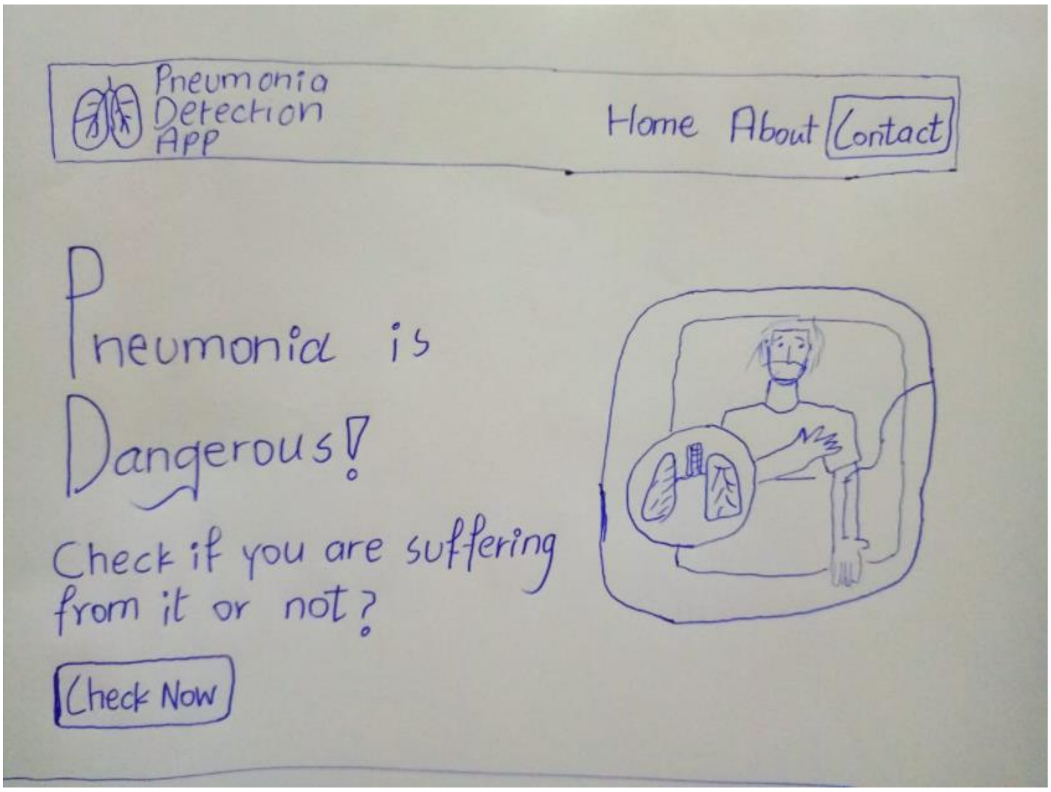


*Figure 14: Low fidelity home page*

**Upload Image Page:**

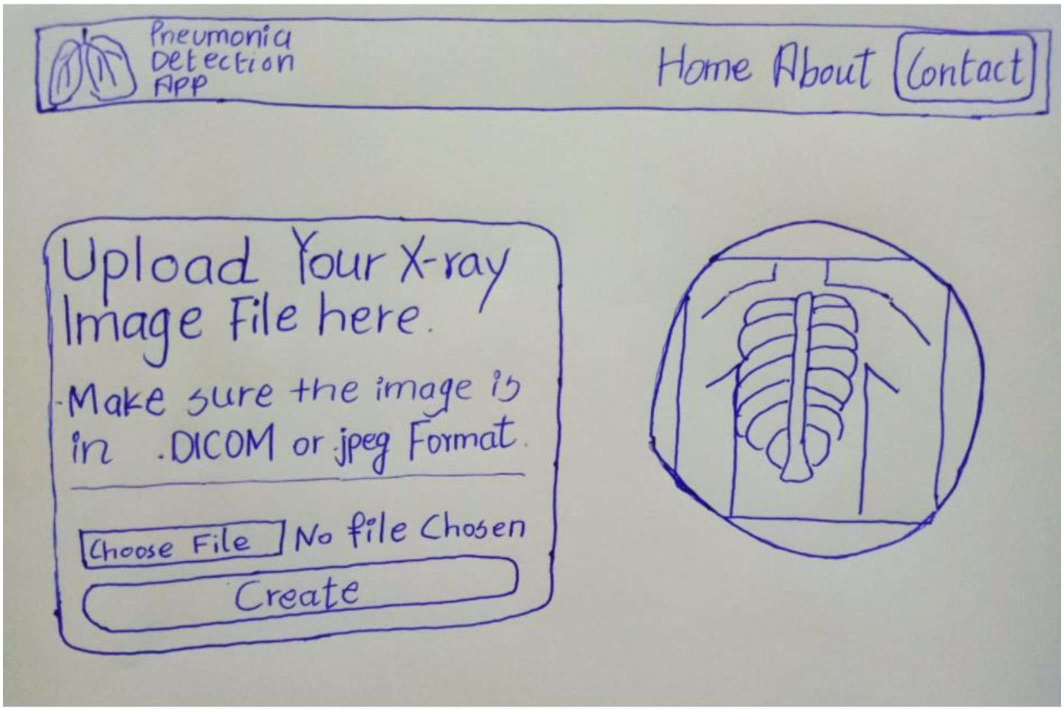

*Figure 15: Low fidelity Upload Image Page*

**Predict Pneumonia Page:**

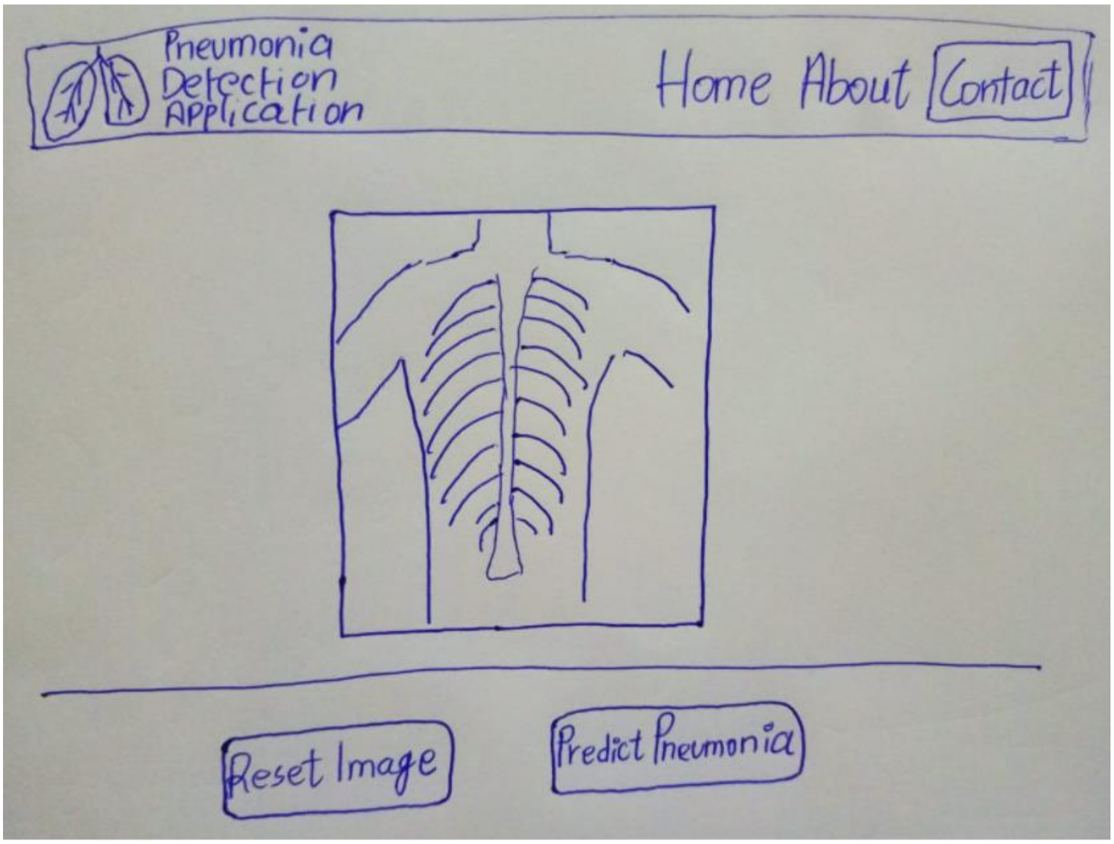


*Figure 16: Low fidelity Predict Pneumonia Page*

**Positive Pneumonia Page:**

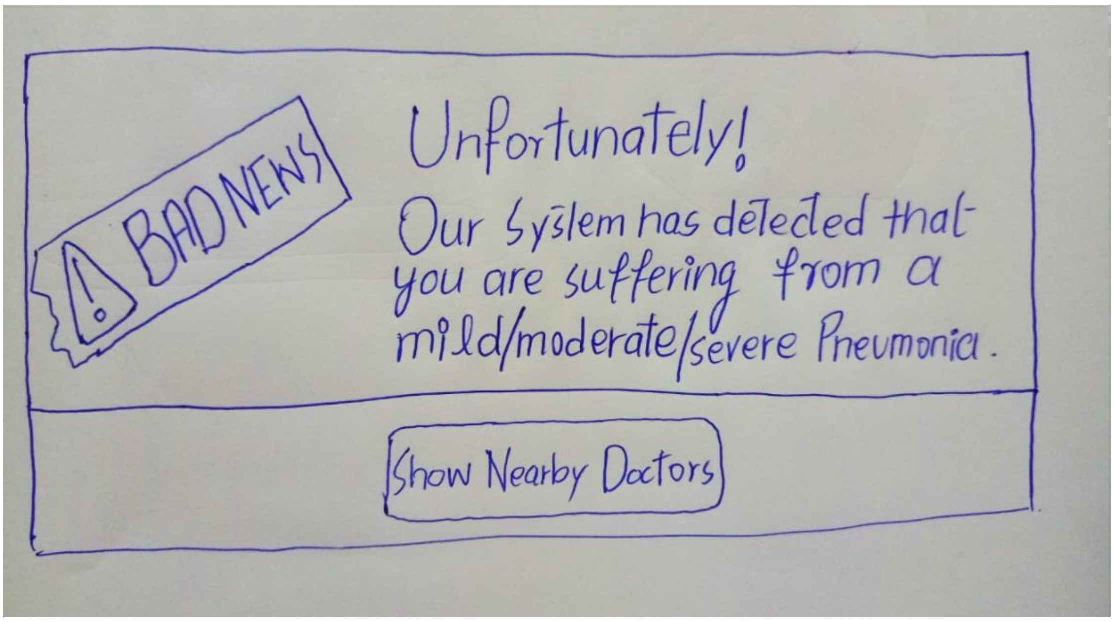


*Figure 17: Low fidelity Positive Pneumonia Page*

**Negative Pneumonia Page:**

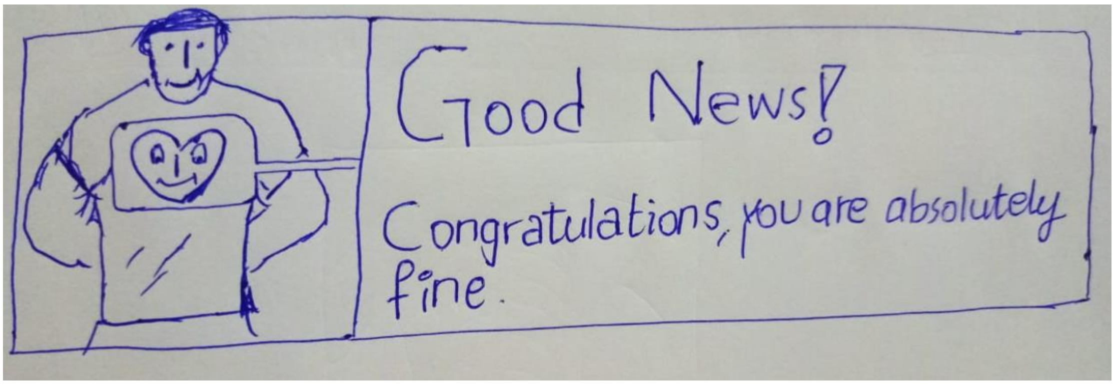


*Figure 18: Low fidelity Negative Pneumonia Page*

**Nearby Doctors Page:**

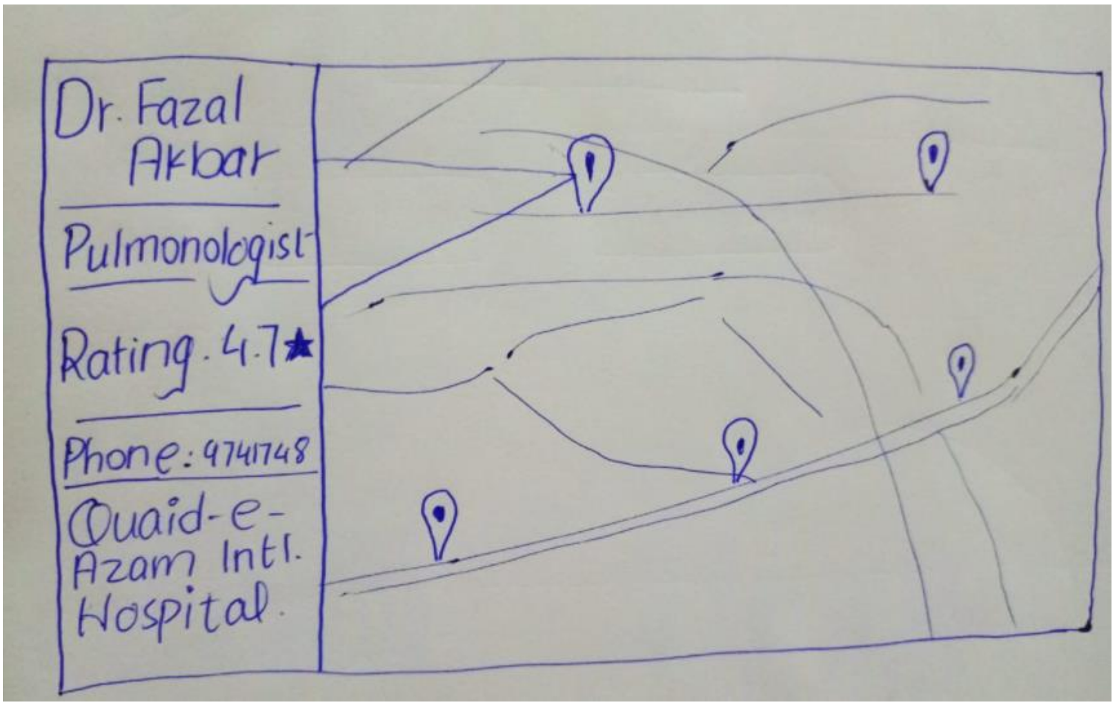


*Figure 19: Low fidelity Nearby Doctors Page*

**Contact Us Page**


*Figure 20: Low fidelity Contact Us Page*

### 4.8.2. High Fidelity Prototype

**Home Page:**

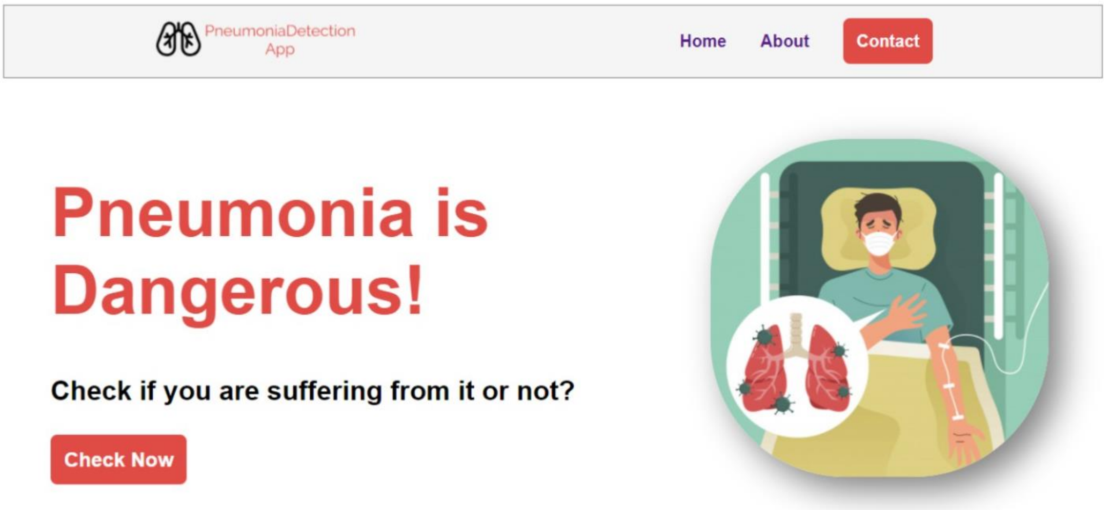


*Figure 21: High fidelity Home Page*

**Upload Image Page:**

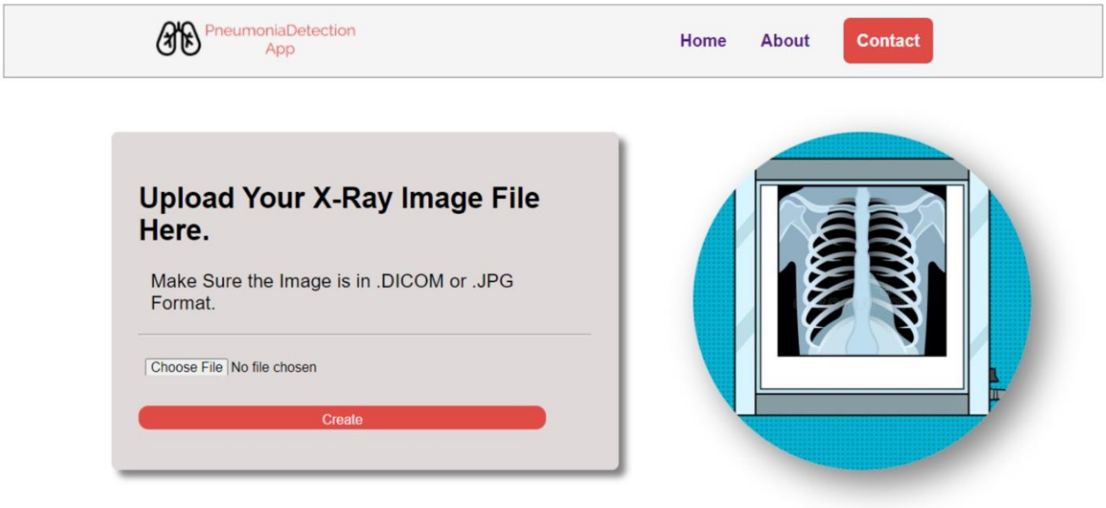


*Figure 22: High Fidelity Upload Image Page*

**Predict Pneumonia Page:**

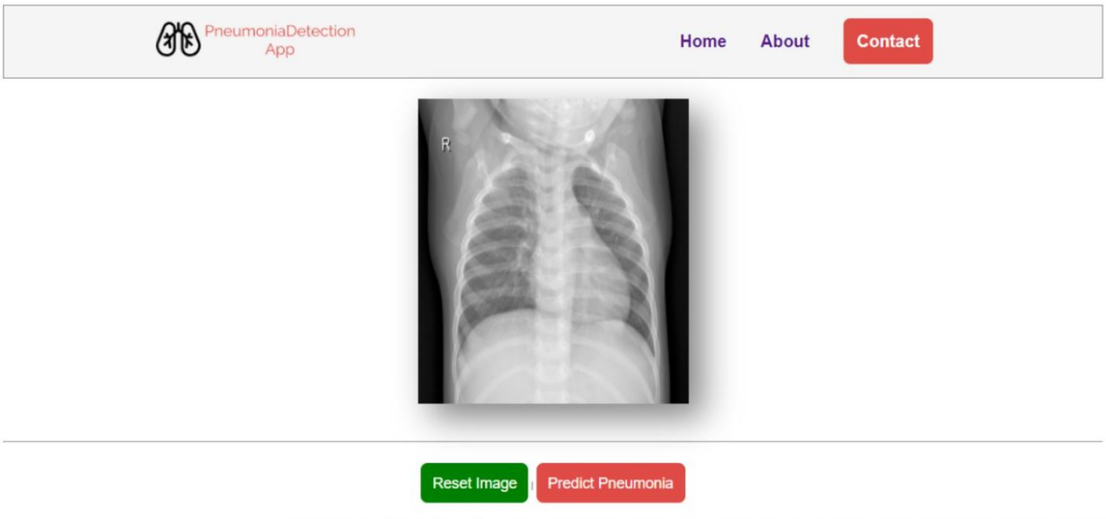


*Figure 23: High Fidelity Predict Pneumonia Page*

**Positive Pneumonia Page:**

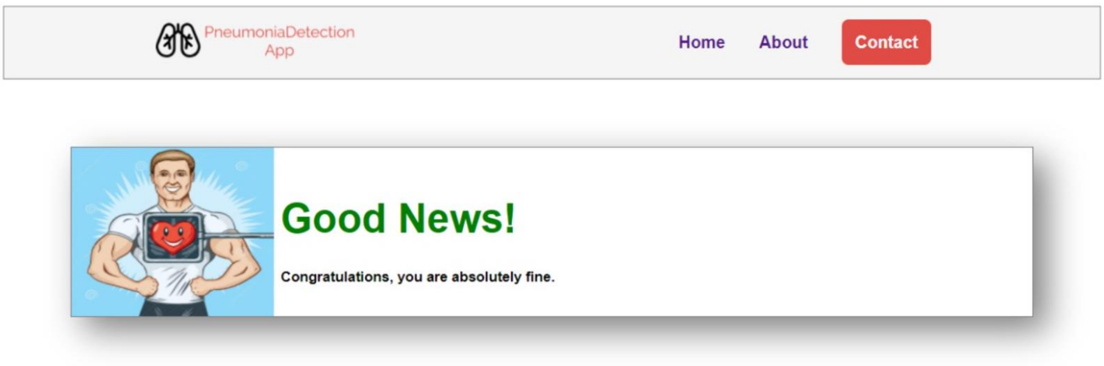


*Figure 24: High Fidelity Positive Pneumonia Page*

## Negative Pneumonia Page:

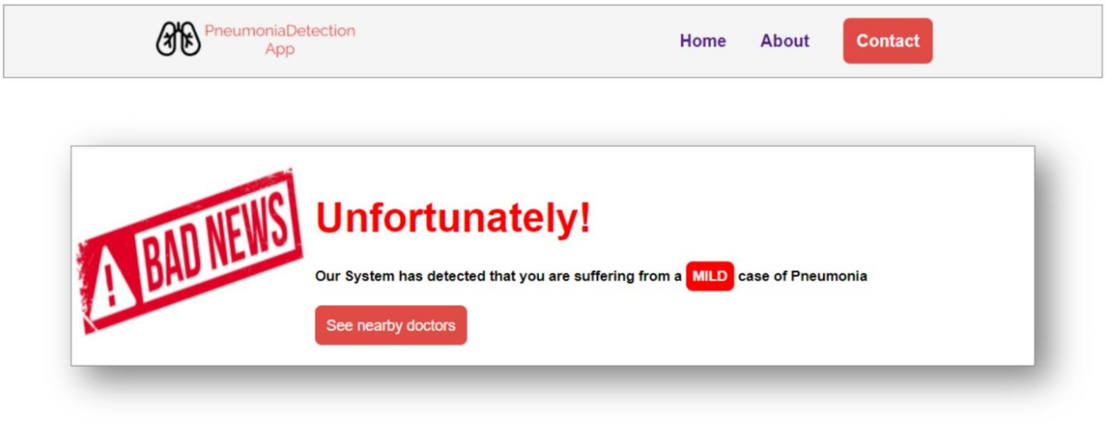


*Figure 25: High Fidelity Negative Pneumonia Page*

## Nearby Doctors Page:

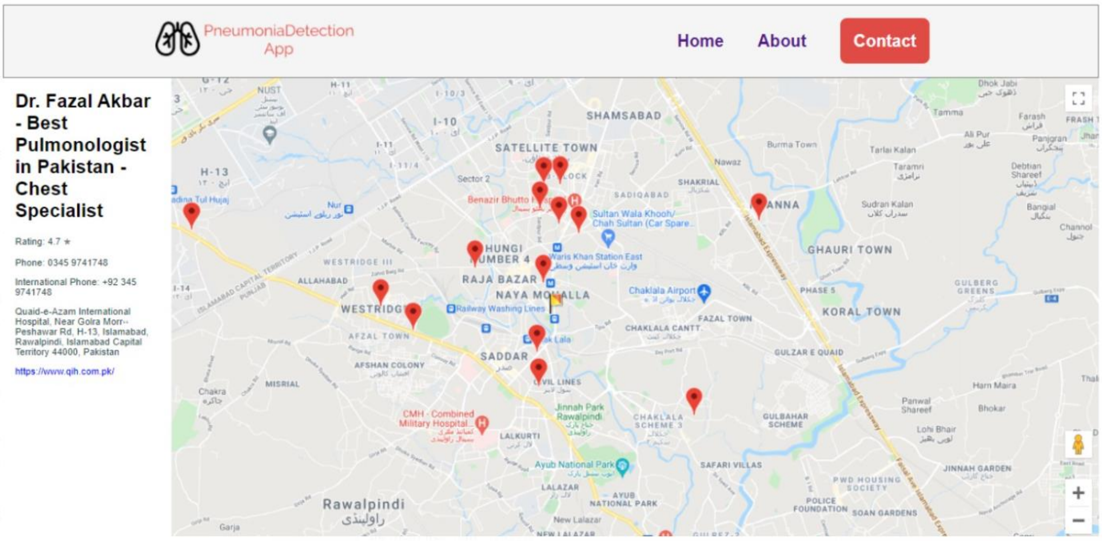


*Figure 26: High Fidelity Nearby Doctors Page*

## Contact Us Page

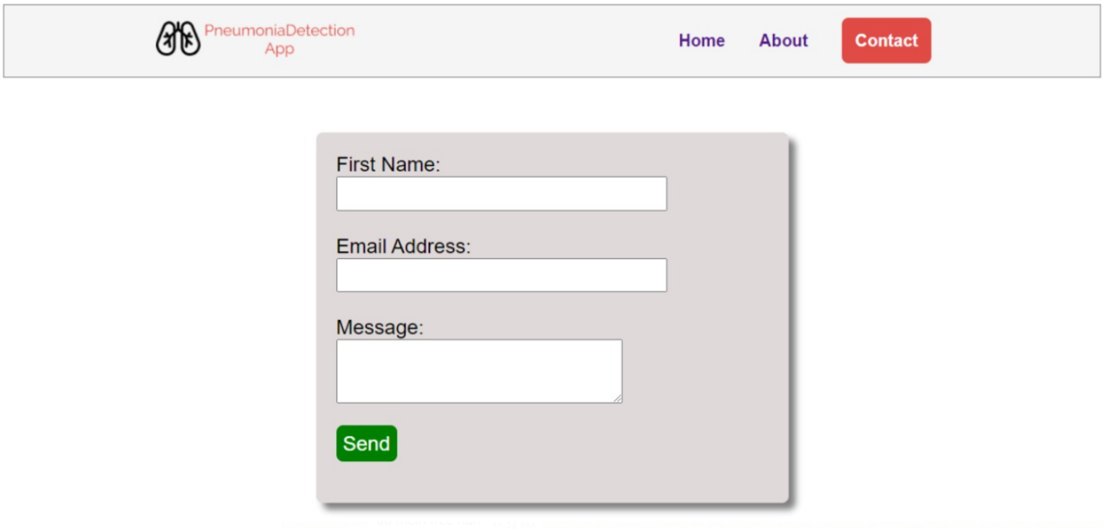


*Figure 27: High Fidelity Contact Us Page*

This chapter has two parts. One gives in-depth detail of system design approach and structure using UML Diagrams. The other part is UI interface that shows how the application will look like and how can users interact with the application

# Chapter 5

# System Implementation

## 5.1. Strategy

The project was divided into two parts. The research where we developed a number of different CNN models using our approach and already established approaches like VGG-16 and inception-v3 etc. and the application development part where we developed a web application and then integrated the two parts.

## 5.2. Tools and Technologies

We used following tools and technologies for this project.

- MS Visual Studio, .NET Core and C# Programming language for backend development of website
- MS Visual Studio Code and Python for CNN model implementation
- Keras and Tensorflow for CNN model Creation
- HTML, CSS and JavaScript for frontend development of website

## 5.3. Development Process

The development approach we adopted for this project is "Incremental Development", as stated earlier in the strategy section, our project is divided into parts or smaller projects mainly research part and development part.

Developing and testing each module/function separately and then integrating them and testing again after integration. We used the Incremental model for the development of this project.

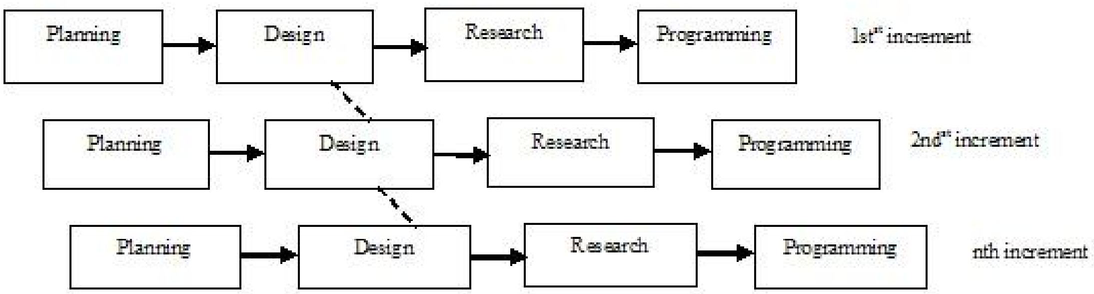


*Figure 28: Development Process*

## 5.4. Material and Methodology

Our CNN model was trained and tested on a public dataset available on kaggle [10]. Our model was developed using Keras, an open-source artificial neural networks interface for Tensorflow library.

### 5.4.1. Dataset

Chest X-ray images were selected from retrospective cohorts of paediatric patients of one to five years old from Guangzhou Women and Children's Medical Centre, Guangzhou. All chest X-ray imaging was performed as part of patients' routine clinical care. The reason for selecting this dataset is that it was tested by two expert physicians before being cleared for AI training.

The original dataset [10] consists of three folders named 'test', 'train' and 'validation'. These folders further contain 2 sub-folders named 'NORMAL' and 'PNEUMONIA'. The dataset contains a total of 5,856 images. These images are divided into test, train and validation as follows:

| Main Folders | Sub Folders | Images |
|---|---|---|
| Test | Normal | 234 |
| | Pneumonia | 390 |
| Train | Normal | 1341 |
| | Pneumonia | 3875 |
| Validation | Normal | 8 |
| | Pneumonia | 8 |

*Table 7: Dataset Division*

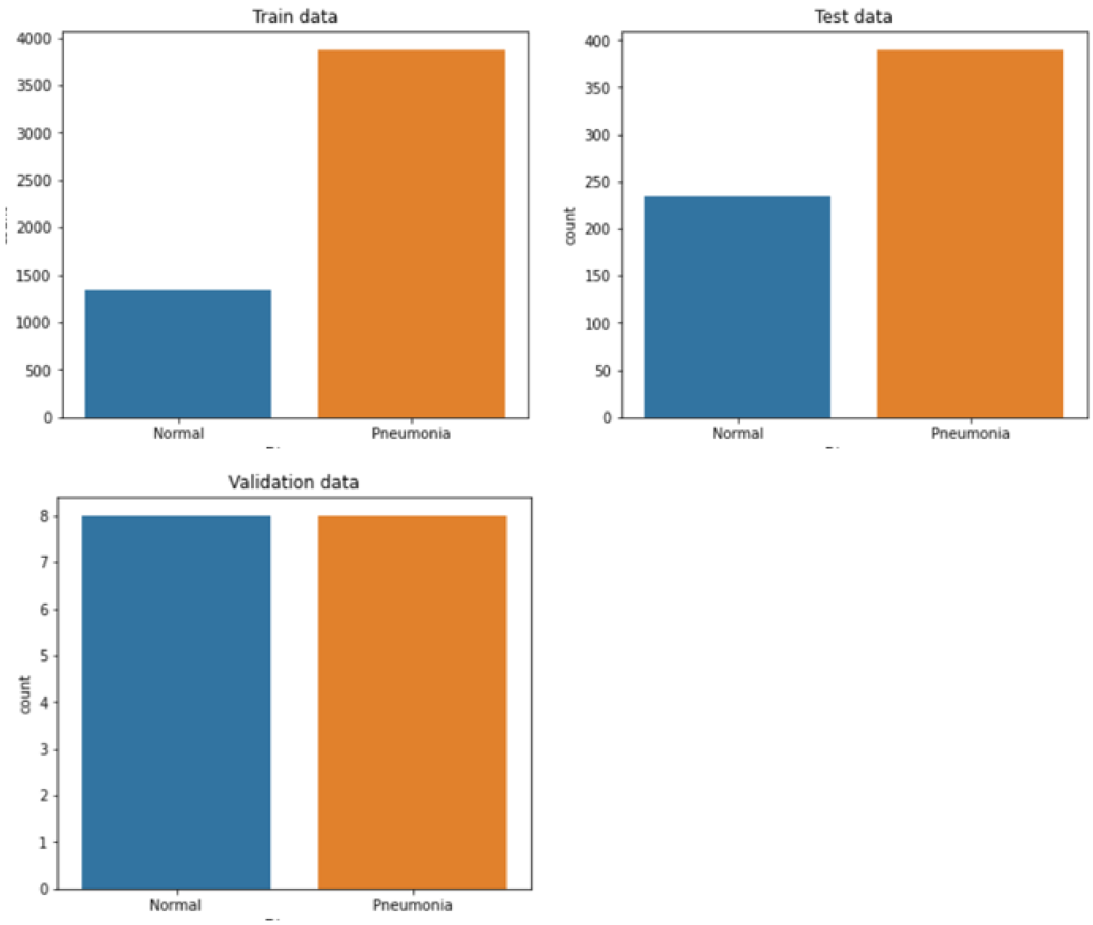


*Figure 29: Test, Train and validation split*

These x-rays are taken in a hospital in China of children aged from 1 to 5 years. Since the validation set was very small, we rearranged the images making 2 folder, 'train' and 'test'.

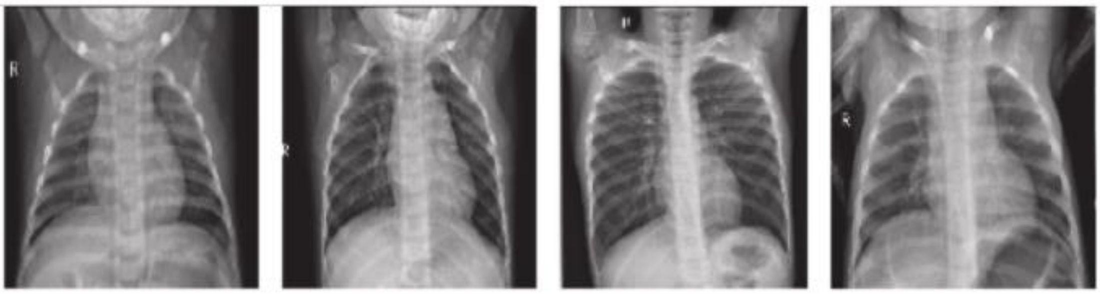

*Figure 30: Sample x-ray normal images*

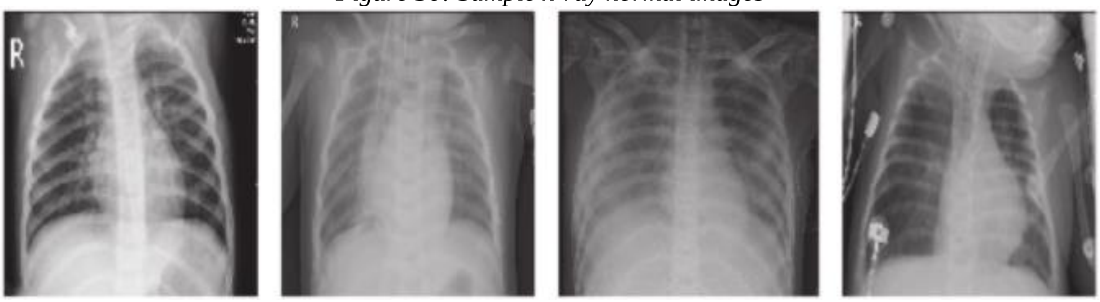

*Figure 31: Sample x-ray pneumonia images*

### 5.4.2. Image Pre-processing and Data Augmentation

Image pre-processing reduces noise from images and enhances the features necessary to predict desired output by the neural network.

For our own developed neural network, we used Data Augmentation techniques in provided by Keras which are as follows:

| Method | Setting |
|---|---|
| Rescale | 1/255 |
| Rotation Range | 30 |
| Shear Range | 0.2 |
| Zoom Range | 0.2 |
| Horizontal Flip | True |
| Vertical Flip | True |

*Table 8: Data Augmentation techniques*

After applying above pre-processing on images, the resulting images are:

*Figure 32: Result of Data Augmentation Applied*

But for the established CNN architectures we used like VGG-16 and inception-v3, we used their own image pre-processing libraries.

## 5.5. CNN Model Implementation

We developed top performance CNN architectures for our problem. VGG-16 and inception-v3 were used to build a CNN models along with our own model. The best performing model was then selected for integration in our application.

### 5.5.1. VGG-16 Architecture:

We used relu optimization function for hidden layers and sigmoid for output layer.

```
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         [(None, 196, 196, 3)]     0

block1_conv1 (Conv2D)        (None, 196, 196, 64)      1792

block1_conv2 (Conv2D)        (None, 196, 196, 64)      36928

block1_pool (MaxPooling2D)   (None, 98, 98, 64)        0

block2_conv1 (Conv2D)        (None, 98, 98, 128)       73856

block2_conv2 (Conv2D)        (None, 98, 98, 128)       147584

block2_pool (MaxPooling2D)   (None, 49, 49, 128)       0

block3_conv1 (Conv2D)        (None, 49, 49, 256)       295168

block3_conv2 (Conv2D)        (None, 49, 49, 256)       590080

block3_conv3 (Conv2D)        (None, 49, 49, 256)       590080

block3_pool (MaxPooling2D)   (None, 24, 24, 256)       0

block4_conv1 (Conv2D)        (None, 24, 24, 512)       1180160

block4_conv2 (Conv2D)        (None, 24, 24, 512)       2359808

block4_conv3 (Conv2D)        (None, 24, 24, 512)       2359808

block4_pool (MaxPooling2D)   (None, 12, 12, 512)       0

block5_conv1 (Conv2D)        (None, 12, 12, 512)       2359808

block5_conv2 (Conv2D)        (None, 12, 12, 512)       2359808

block5_conv3 (Conv2D)        (None, 12, 12, 512)       2359808

block5_pool (MaxPooling2D)   (None, 6, 6, 512)         0

flatten (Flatten)            (None, 18432)             0

dense (Dense)                (None, 2)                 36866
=================================================================
Total params: 14,751,554
Trainable params: 36,866
Non-trainable params: 14,714,688
```

*Figure 33: VGG-16 Architecture*

### 5.5.2. Inception v-3

Inception-v3 uses RMSProp Optimization function for hidden layers and sigmoid for output layer

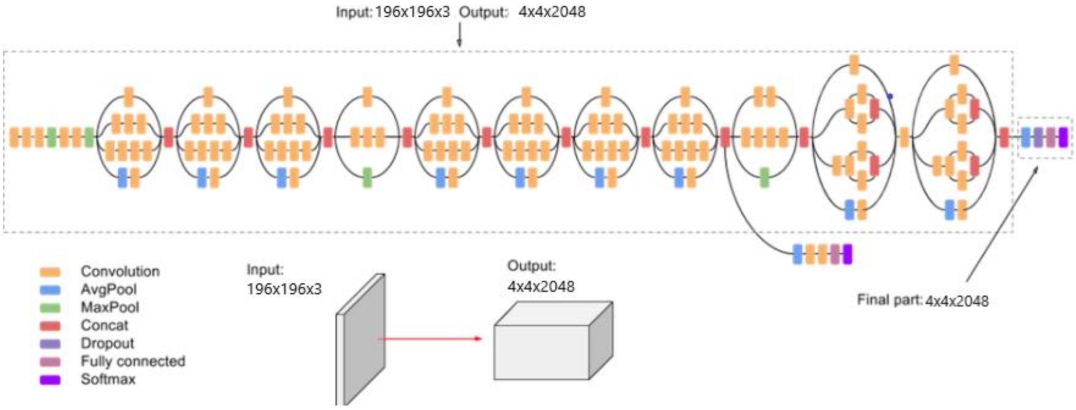


*Figure 34: Inception Architecture*

### 5.5.3. Models Selection

A total of three models were trained and then tested. The best performing model was then selected to implement and integrate with the application. Following are the performances of the three models.

| Model | Testing/Validation Accuracy |
|---|---|
| Inception-v3 | 98.8 |
| VGG-16 | 97.9 |
| Self-developed | 90.9 |

*Table 9: CNN Models Results*

As we can see from the table that inception-v3 is the best performing model among all three. Inception is created by google and has the highest accuracy from all other established CNN architectures. Since inception-v3 has the highest accuracy, therefor we used it in our application for pneumonia detection.

In this chapter we have discussed system implementation and gave details of CNN model creation, training and testing. Furthermore we set a criteria for finalizing the model to be used in our application. Inception-v3 had the highest accuracy on test set, therefor we selected inception-v3.

# Chapter 6

# System Testing & Evaluation

## 6.1. Test Strategy

## 6.2. Component Testing

Component testing can be referred to as Unit Testing or Module Testing. The first step and the important step is testing of the components before integration to make sure each part and each component is working fine. At first, we created the CNN model in Python and then trained it with Kaggle Dataset. After that, we tested that model using different X-Ray Files and it worked fine with more than 90% accuracy. Once we got satisfied with the model, we then developed a website in ASP.net core, designed all the pages and then integrated our model into it.

Then we tested Google Maps API and tried to fetch nearby doctors and it worked fine. Initially we had problems regarding locations, but eventually we solved that issue and then it worked fine.

## 6.3. Unit Testing

After Component testing, we moved on to Unit Testing. Unit testing involves testing each unit of the system separately. We distributed our system into three units and then tested each one of them separately to make sure no errors remains.

We first tested Image Uploading unit using .jpeg and .dicom X-Ray Files. Initially we had some problems regarding .dicom uploading, but after some tweaks, we solved it and finally the unit testing for this unit was successful.

Then we tested Pneumonia prediction unit in which we tested various image files for prediction and the testing was successful after several attempts.

Finally we tested Google Maps Locations unit, in which tested if the location of the user is fetched accurately and all the nearby doctors are shown or not? The test was successful after several attempts.

## 6.4. Integrated Testing

After the unit testing, we move to the Integration testing. It is a level of software testing where we have to test the components after integration. We have tested the system and the interaction of the integrated units.

We took all the components, and then performed integrated testing by integrating different components with each other.

First, we integrated image uploading with our website and it went fine without any issues. Then we integrated CNN model with our website and performed Pneumonia prediction, it went fine, and testing was successful.

Finally we integrated User location and Google Maps API to show nearby doctors and the testing was successful.

### 6.5. System Testing

After integration, we again tested the whole system, initially we had to face a lot of errors because after several amendments in the code, we finally made it and the whole system ran successfully without any error. We tested the website multiple times and it worked fine without any errors.

### 6.6. Test Cases

Although there are several vague to complex test cases in this application, we have added only those that are directly related to the operational use of the application.

#### 6.6.1. Test Case#1 (User Uploads his X-Ray Image File)

| | |
|---|---|
| Preconditions | The web application must be opened and connected to the Internet Connection. |
| Actions | User opens the web application and opens the upload image page.<br>User clicks on Select Image Button and selects his X-Ray File.<br>User clicks on Upload Button. |
| Expected Results | If the format of the Image File is either .jpeg or .dicom, the image will be uploaded.<br>Else the image will not be uploaded. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | High |
| Input Data | Click on Upload Image. |

*Table 10: Test Case 1*

### 6.6.2. Test Case#2 (User gives the command for Pneumonia Prediction)

| | |
|---|---|
| Preconditions | The X-ray Image File must be uploaded before prediction. |
| Actions | User uploads his X-ray image file.<br>The image is uploaded and shown to him on a separate page with predict option.<br>User clicks on Predict button. |
| Expected Results | The predict command will be sent to the system.<br>The CNN model will analyse the image and provide result.<br>The prediction result will be shown to the user on a new page.<br>Either the Pneumonia result will be positive or negative.<br>If the Pneumonia result is positive, it is either mild, Moderate, or severe. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | High |
| Input Data | Click on Predict Pneumonia. |

*Table 11: Test Case 2*

### 6.6.3. Test Case #3 (The system predicts Pneumonia)

| | |
|---|---|
| Preconditions | The user must upload his X-Ray Image File and give "Predict Pneumonia" command. |
| Actions | The system takes the X-Ray Image File from the user.<br>The system runs it through CNN model.<br>The system takes the result from CNN model and gives the result back to the user. |
| Expected Results | The system takes the Pneumonia result from the CNN model and show it to the user. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | High |
| Input Data | The user gives Predict command. |

*Table 12: Test Case 3*

### 6.6.4. Test case #4 (User asks for nearby doctors)

| | |
|---|---|
| Preconditions | The Pneumonia result for the patient/user must be positive for this option to be shown. |
| Actions | The user clicks on "Show nearby doctors" button.<br>The system predicts that the patient has pneumonia.<br>The system shows the option to check nearby doctors. |
| Expected Results | The Google maps will open showing all those doctors (pulmonologists) that have clinics near the patient's home. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | Medium. |
| Input Data | Click on "Show nearby doctors". |

*Table 13: Test Case 4*

### 6.6.5. Test case #5 (Google Maps fetches user location)

| | |
|---|---|
| Preconditions | The user must give the system the permission for the location to be fetched before this. |
| Actions | The Google Maps API is called and system asks user for location permission.<br>If the user gives the permission, the location is fetched. |
| Expected Results | If the user gives the permission, the system fetches the location.<br>If the user denies the permission, the system does not fetches the location. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | High |
| Input Data | The user clicks on “Show nearby doctors”. |

*Table 14: Test Case 5*

### 6.6.6. Test case #6 (Google Maps show nearby Doctors)

| | |
|---|---|
| Preconditions | The system must fetch user location before this. |
| Actions | The user clicks on “show nearby doctors” and the Google Maps API is called.<br>Google Maps fetches the location of the user.<br>Google Maps loads nearby doctors. |
| Expected Results | A google Map will be shown with all the nearby doctors shown prominently along with their details. |
| Tested by | Muhammad Basit Ali and Uzair Abdullah Mir. |
| Result | Pass |
| Priority | High |
| Input Data | Click on “Show nearby doctors”. |

*Table 15: Test Case 1*

### 6.7. Results & Evaluation

We can see the results obtained in the above test cases that every test has passed successfully. Any test that failed was corrected until all the tests passed.

In this chapter we have discussed the testing done on the system. The testing was done throughout the development of the project.

First we did unit testing by testing each functionality and module after its development before they were all integrated. Unit testing was performed informally. On successful unit testing, we moved on integration and system testing. These were carried out successfully and all the test cases passed.

# Chapter 7

# Conclusion

## 7.1. Contributions

Research work done on detection of diseases from medical imaging is at peak. More and more individuals, students and researches and even giant techs are investing their time and resources and we will hopefully see some breakthrough coming years that will change the healthcare industry as we know it for good.

Our contribution in the medical domain is unique in a sense that we have contributed in terms of research and development. Although there has been research published on different medical imaging, but no one has yet developed an application to put that research to practical. We have developed an (web) application that will allow people to get diagnosed from anywhere through internet.

## 7.2. Reflections

Looking back at our project, we think that it is a very different project because there are no applications that provide such facilities. The reason maybe because 100 percent accuracy has not yet being achieved but looking at the research being carried out in this domain, we think we can see some great inventions in the future.

Our application has great benefits on the society especially in this pandemic where hospitals are not as safe due to covid. If developed further, it can change the way people get diagnosed and treated.

## 7.3. Future work

Covid-19 has changed the way we used to carry out daily life activities from education to work and healthcare. It is getting hard to get appointments because of limited availability of doctors. NESCOM Hospital in H-11 has limited the number of visitors by limiting the availability of doctors to avoid crowding in hospital. Therefor they have been working of going online. Our project has great potential if further developed. Some of the modules/ functionalities that can be further implemented are:

- Patients could get an appointment and have online video call with the doctors.
- Other pulmonary diseases detection can be implemented as in [9] to further extend the area of automated detection of diseases.

This could give us an online health care system that can greatly reduce the need to visit the hospitals and limit the spread if viruses especially Covid-19.

# REFERENCES

Add your reference here in IEEE citation style.


[1] "One child dies of pneumonia every 39 seconds, agencies warn", Unicef.org, 2021. [Online]. Available: https://www.unicef.org/pakistan/press-releases/one-child-dies-pneumonia-every-39-seconds-agencies-warn. [Accessed: 01- Jun-2021].

[2] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, p. 2402, 2016.

[3] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[4] A. Y. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," Nat. Med., vol. 25, no. 1, pp. 65–69, 2019.

[5] M. Grewal, M. M. Srivastava, P. Kumar, and S. Varadarajan, "RADNET: Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans," arXiv [cs.CV], 2017.

[6] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks," Radiology, vol. 284, no. 2, pp. 574–582, 2017.

[7] P. Huang et al., "Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: A matched case-control study," Radiology, vol. 286, no. 1, pp. 286–295, 2018.

[8] K. E. Asnaoui, Y. Chawki, and A. Idri, "Automated methods for detection and classification pneumonia based on X-Ray images using deep learning," arXiv [eess.IV], 2020.

[9] P. Rajpurkar et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv [cs.CV], 2017.

[10] P. Mooney, "Chest X-ray images (pneumonia)." .

FYP

ORIGINALITY REPORT

| 14% | 6% | 2% | 11% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to Higher Education Commission Pakistan<br>Student Paper | 2% |
| 2 | Submitted to Edge Hill University<br>Student Paper | 1% |
| 3 | Submitted to Bournemouth University<br>Student Paper | 1% |
| 4 | www.se.rit.edu<br>Internet Source | 1% |
| 5 | Submitted to Gulf College Oman<br>Student Paper | 1% |
| 6 | Submitted to Multimedia University<br>Student Paper | 1% |
| 7 | Submitted to Chester College of Higher Education<br>Student Paper | <1% |
| 8 | Submitted to University of Paisley<br>Student Paper | <1% |
| 9 | Submitted to University of Surrey | |

| | | |
|---|---|---|
| | Student Paper | <1% |
| 10 | Submitted to University of Wolverhampton<br>Student Paper | <1% |
| 11 | www.coursehero.com<br>Internet Source | <1% |
| 12 | www.slideshare.net<br>Internet Source | <1% |
| 13 | Submitted to Universiti Teknologi Malaysia<br>Student Paper | <1% |
| 14 | Submitted to Informatics Education Limited<br>Student Paper | <1% |
| 15 | Submitted to Thames Valley University<br>Student Paper | <1% |
| 16 | Submitted to Cork Institute of Technology<br>Student Paper | <1% |
| 17 | Submitted to University of Greenwich<br>Student Paper | <1% |
| 18 | arxiv.org<br>Internet Source | <1% |
| 19 | Submitted to SASTRA University<br>Student Paper | <1% |
| 20 | Submitted to Southampton Solent University<br>Student Paper | <1% |

| | | | |
|---|---|---|---|
| 21 | Submitted to University of Birmingham<br>Student Paper | <1% |
| 22 | Submitted to CSU, San Jose State University<br>Student Paper | <1% |
| 23 | Submitted to Taibah University<br>Student Paper | <1% |
| 24 | Submitted to University of Keele<br>Student Paper | <1% |
| 25 | Submitted to Queen Mary and Westfield College<br>Student Paper | <1% |
| 26 | Submitted to University of Malaya<br>Student Paper | <1% |
| 27 | Submitted to University of Westminster<br>Student Paper | <1% |
| 28 | Wasif Khan, Nazar Zaki, Luqman Ali. "Intelligent Pneumonia Identification from Chest X-Rays: A Systematic Literature Review", Cold Spring Harbor Laboratory, 2020<br>Publication | <1% |
| 29 | murray.newcastle.edu.au<br>Internet Source | <1% |
| 30 | Submitted to Bahrain Polytechnic<br>Student Paper | <1% |

rcciit.org

| | | |
|---|---|---|
| 31 | Internet Source | <1% |
| 32 | uniassignment.com<br>Internet Source | <1% |
| 33 | www.kaggle.com<br>Internet Source | <1% |
| 34 | www.scribd.com<br>Internet Source | <1% |
| 35 | Anwaar Ulhaq, Jannis Born, Asim Khan, Douglas Pinto Sampaio Gomes, Subrata Chakraborty, Manoranjan Paul. "COVID-19 Control by Computer Vision Approaches: A Survey", IEEE Access, 2020<br>Publication | <1% |
| 36 | Karan Jakhar, Nishtha Hooda. "Big Data Deep Learning Framework using Keras: A Case Study of Pneumonia Prediction", 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018<br>Publication | <1% |
| 37 | home.eps.hw.ac.uk<br>Internet Source | <1% |
| 38 | www.researchgate.net<br>Internet Source | <1% |

| | | |
|---|---|---|
| 39 | Okeke Stephen, Mangal Sain, Uchenna Joseph Maduh, Do-Un Jeong. "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare", Journal of Healthcare Engineering, 2019<br>Publication | <1% |
| 40 | senior.ceng.metu.edu.tr<br>Internet Source | <1% |
| 41 | www.unicef.org<br>Internet Source | <1% |

| | | | |
|---|---|---|---|
| Exclude quotes | On | Exclude matches | Off |
| Exclude bibliography | On | | |