PNEUMONIA DETECTION USING AI

Submitted in partial fulfillment of the requirements

of the degree of

Bachelor of Engineering in Information Technology

By

OMKAR WADEKAR (18101A0048)
RITESH RAHATAL (18101A0052)
MANASI SHERKHANE (18101A0062)

Under the Guidance of

Prof. NEHA KUDU

Department of Information Technology

Vidyalankar Institute of Technology

Wadala(E), Mumbai-400437

University of Mumbai

2021-22

CERTIFICATE OF APPROVAL

This is to certify that the project entitled
"PNEUMONIA DETECTION USING AI"
is a bonafide work of

OMKAR WADEKAR (18101A0048)
RITESH RAHATAL (18101A0052)
MANASI SHERKHANE (18101A0062)

submitted to the University of Mumbai in partial fulfillment of the requirement for the

award of the degree of

Undergraduate in **"INFORMATION TECHNOLOGY"**.

Guide	Head of Department	Principal
(Prof. Neha Kudu)	(Dr. Vipul Dalal)	(Dr. S.A. Patekar)

Project Report Approval for B. E.

This project report entitled ***Pneumonia Detection Using AI*** by

1. ***Omkar Wadekar (18101A0048)***
2. ***Ritesh Rahatal (18101A0052)***
3. ***Manasi Sherkhane (18101A0062)***

is approved for the degree of ***Bachelor of Engineering in Information Technology.***

Examiners

1. Karshali 25/04/22 Karshali Bhusnad -

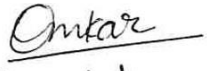
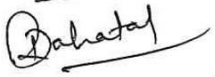
2. NRkudu 25/04/22 Neha Kudu -

Date: 20/04/22

Place: Mumbai

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/sourcein our submission. We understand that any violation of the above will be cause for disciplinaryaction by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

	Name of student	Roll No.	Signature
1.	Omkar Wadekar	18101A0048	
2.	Ritesh Rahatal	18101A0052	
3.	Manasi Sherkhane	18101A0062	

Date: 20-04-2022

ACKNOWLEDGEMENT

We are honored to present "**Pneumonia Detection Using AI**" as our B.E Final Year Project. We are using this opportunity to express our profound gratitude to our principal "**Dr. Sunil Patekar**" for providing us with all proper facilities and support.

We express our deepest gratitude towards our HOD "**Dr. Vipul Dalal**" for his valuable and timely advice during the various phases in our project. We would like to thank our project guide **"Prof. Neha Kudu"** for support, patience and faith in our capabilities and for giving us flexibility in terms of working and reporting schedules. Finally, we would like to thank everyone who has helped us directly or indirectly in our project.

We would also like to thank our staff members and lab assistant for permitting us to use computers in the lab as when required. We thank our college for providing us with excellent facilities that helped us to complete and present this project.

Omkar Wadekar

Ritesh Rahatal

Manasi Sherkhane

ABSTRACT

Pneumonia is known as one of the most life-threatening diseases in the world which affects the lung(s) of humans and is known as one of the leading causes of death in India. Roughly 33% of the deaths in India are caused as a result of pneumonia and as reported by the World Health Organization (WHO). Currently pneumonia is diagnosed using a Chest X-Ray image which is then evaluated by an expert radiotherapist. This process is quite exerting and travail and it often leads to a difference in opinion among the experts. Thus, developing a solitary system would bebeneficial for identification, preventing further transmission and treatment in remote areas.

This system proposes a CNN model which has been trained from scratch and that will classify and also detect the presence of pneumonia from a dataset of chest X-ray images. For this system the CNN model would extract features from a given dataset of chest X-ray image and then classify it to work it out if an individual is infected with pneumonia or not.

CONTENTS

SR NO.	TOPIC	PAGE NO.
	ABSTRACT	v
	LIST OF FIGURES	vii
1	**INTRODUCTION** 1.1 Introduction 1.2 Scope 1.3 Motivation 1.4 Problem Statement	1 2 3 3 4
2	**LITERATURE SURVEY** 2.1 Survey of Existing or Similar System 2.2 Limitation Existing/Similar System or Research gap	5 6 8
3	**SYSTEM DESIGN** 3.1 Proposed solution 3.2 Features of Pneumonia Detection System 3.3 Algorithm and Process Design 3.4 Data Flow Diagram	9 10 10 11 12
4	**ANALYSIS** 4.1 Methodology and Implementation 4.2 Feasibility Study 4.3 Gantt chart	13 14 16 17
5	**DESIGN** 5.1 UML Diagram 5.1.1 Use Case Diagram 5.1.2 Activity Diagram 5.1.3 Sequence Diagram 5.2 Hardware and software requirements	18 19 19 20 21 22
6	**RESULTS AND DISCUSSION**	23
7	**CONCLUSION AND FUTURE SCOPE** 7.1 Conclusion 7.2 Future scope	28 29 29
	REFERENCES	30

GitHub link - https://github.com/riteshrahatal/Pneumonia-Detection-Using-AI

LIST OF FIGURES

Figure No.	Figure Name	Page No.
3.4	Data Flow Diagram – Level 0	12
5.1.1	Use Case Diagram	19
5.1.3	Sequence Diagram	21
5.1.2	Activity Diagram	20
4.3	Gantt Chart	17
6.1	Home Page	24
6.2	Uploaded Image File	25
6.3	Final Output(Normal)	26
6.4	Final Output(Pneumonia)	27

CHAPTER 1
INTRODUCTION

INTRODUCTION

1.1 Introduction

One of the most important diseases prevalent all across the globe is pneumonia, which affects the lungs. Pneumonia is a deadly disease that causes the air sacs in the lungs to become filled with pus and liquid. Pneumonia is divided into two important types, which are bacterial type and viral type. Pneumonia is a disease that disturbs the alveoli of the lungs and has a mortality account which accounts for about 16% of the world deaths [1], being the world's leading cause for deaths among many children.

It is an acute respiratory infection of the lungs, and it has a high incidence, accounting for about 12% of the total population. Nowadays, the incidence of pneumonia is still increasing due to the social population aging, increased immune-impaired hosts, pathogen changes, difficult pathogenic diagnosis, and increased bacterial resistance. A chest radiograph analysis (CXR) is the most commonly used method for X-ray examination to diagnose and differentiate the type of pneumonia. However, because of a lack of professional radiologists, pneumonia has alarming mortality rates in some limited resource areas. Therefore, addressing the issue of how to improve the accuracy of pneumonia detection and reducing the cost of pneumonia detection has great help for the treatment and prevention of pneumonia.

In recent years, deep learning technologies have developed rapidly. Deep learning is a widely used tool in research fields such as computer vision, speech analysis, and natural language processing. This method is particularly suitable for those fields that need to analyze large amounts of data and human intelligence. A major advantage of using deep learning methods is that complex features can be learned directly from the raw data. This allows us to define a system that does not rely on manual operations, which is unique among other machine learning technologies. The use of deep learning as a machine learning and pattern recognition tool is also becoming an important aspect in the field of medical image analysis. At present, the deep learning technology has played an important role in medical image processing, computer-aided diagnosis, image interpretation,

image fusion, image registration, image segmentation, and image-guided therapy. It can help doctors diagnose and predict disease risk accurately and quickly.

1.2 Scope

For improvements in the project we can adapt to other convolutional neural network architectures like Inception-v3, shuffle Net, and Mobile Net architectures for pneumonia detection can be implemented and the optimization of hyper-parameters should also be considered to improve the accuracy of the model.Also, the performance of the proposed Deep CNN model can be improved with more number of layers and parameters.

This project can help medical workforce in their decision making for a real-time application of the use of accurate models in detecting pneumonia and discover the potential of diagnosing pneumonia using deep learning. Different lung disease classes can also be included in this model to detect various lung diseases using the chest X-ray images.This will allow clinicians to recognize lung diseases from chest X-ray images with lower prevalence at an earlier stage of the disease.

1.3 Motivation

Pneumonia affects a large number of individuals, especially children, mostly in developing and underdeveloped countries characterized by risk factors such as overcrowding, poor hygienic conditions, and malnutrition, coupled with the unavailability of appropriate medical facilities. Early diagnosis of pneumonia is crucial to cure the disease completely. Examination of X-ray scans is the most common means of diagnosis, but it depends on the interpretative ability of the radiologist and frequently is not agreed upon by the radiologists. Thus, an automatic CAD system with generalizing capability is required to diagnose the disease.

To the best of our knowledge, most previous methods in the literature focused on developing a single CNN model for the classification of pneumonia cases, and the use of the ensemble learning

paradigm in this classification task has not been explored. However, the ensemble learning model incorporates the discriminative information from all the constituent base learners, allowing it to make superior predictions, and thus was implemented in this study. To handle the low amount of available biomedical data, transfer learning models were used as base learners, the decision scores of which were ensembled.

1.4 Problem Statement

Properly diagnosing pneumonia can be a tall order because it requires the review of chest radiographs (CXR) by highly trained specialists. The specialist confirms a case of pneumonia by also examining the patient's clinical history, vital signs and laboratory examination results. Pneumonia usually reveals itself in the lungs as an area(s) of increased opacity on CXR, however, diagnosis on CXR can be complicated by several other lung conditions. These include fluid overload (pulmonary edema), bleeding, volume loss (atelectasis), lung cancer or post-radiation/surgical changes. Outside of the lungs, fluid in the pleural space (pleural effusion) also appears as increased opacity on CXR.

CXRs are the most commonly performed diagnostic imaging study. Several factors such as positioning of the patient and depth of inspiration can alter the appearance of the CXR, complicating interpretation further. In addition, clinicians are faced with reading high volumes of images every shift. This project will attempt to develop a model that will detect visual signals of pneumonia in medical images. It will automatically locate lung opacities on CXRs.

CHAPTER 2
LITERATURE SURVEY

LITERATURE SURVEY

2.1 Survey of Existing or Similar System

The latest improvements in the field of Machine Learning and AI mainly due to large scale usage of Convolutional Neural Networks (CNNs) and the availability of free dataset. That was once considered to be very rare and has assisted various algorithms to perform much better than was not considered to be a commonplace a few years ago. The automated diagnosis of varied diseases has received a high interest. The low performance of several CNN models on diverse abnormalities proves that a single model cannot be used for all the purposes. So for a better exploration of machine learning in the chest screening, Wang et al. (2017) [2] released a larger dataset of frontal chest X-Rays.

Huang et al. [3] adopted deep learning techniques. Performance of different variants of Convolutional Neural Networks for abnormality detection in chest X-Rays was then proposed by Islam et al. using the publicly available OpenI dataset for the better exploration of machine learning in chest screening.

Cicero et al. discussed the training and validation of CNNs with modest-sized medical data to detect pathology in 2017 [4]. Ma et al. presented a survey on deep learning for pulmonary medicalimaging in 2019 [5]. Jaiswal et al. from University of Bedfordshire described an approach based on deep learning for identifying pneumonia in chest X-way in 2019 [6].

In recent times, Pranav Rajpurkar, Jeremy Irvin, et al. (2017) [7] explored the dataset for detecting pneumonia at a level far better than radiologists, they referred their model as ChexNet that uses DenseNet-121 layer architecture for detecting all the 14 diseases from a lot of 112,200 images that are available in the dataset.

After the CheXNet [7] model, Benjamin Antin et al. (2017) [8] worked on the same dataset and proposed a logistic regression model for detecting pneumonia.

Okeke Stephen, Mangal Sain, Uchenna, and Do-Un Jeong In [9] proposed a CNN model **t**hat's trained from scratch to classify and to detect the presence of pneumonia from a collection of chest X-ray image dataset. Unlike other methods that rely only on transfer learning approaches to realize an interesting classification performance, they constructed a convolutional neural network model from scratch that extracts features from a given chest X-ray image and classifies it to work out if an individual is infected with pneumonia. This model could help mitigate the reliability and interpretability challenges often faced when handling medical imagery.

2.2 Limitation Existing/Similar System or Research gap

Following limitations were observed during the development of the model:

1. Processing and building the model requires fast and efficient processors which is time and cost consuming.
2. 100% accuracy is not achievable.
3. The challenges of large variation in sensing modality which is complicated by human anatomy are faced in medical image analysis.
4. Parameters affecting medical images fluctuate from organ to organ.
5. Medical Images are affected by noise due to sensors, device implantation, electronics leading to inefficiency while detection.

CHAPTER 3
SYSTEM DESIGN

SYSTEM DESIGN

3.1 Proposed Solution

This study proposes a convolutional neural network model trained from scratch to classify and detect the presence of pneumonia from a collection of chest X-ray image samples. We have used the x-ray images data in our model. We have constructed our own convolutional neural network model from scratch to extract features from a given chest X-ray image and classify it to determine if a person is infected with pneumonia.

The pneumonia affected parts of the lung would be detected using computer vision. An optimum solution for the detection of pneumonia from chest X-rays is proposed in this project.

3.2 Features of Pneumonia Detection System

1. The web portal is used to detect Pneumonia by uploading an X-ray .jpeg image on the website and within seconds the results are displayed with an accuracy of 90%.
2. The patient can use this application by sitting at home in their own comfort.
3. The application of clinical image diagnosis of pneumonia X-rays can reduce the workload of clinicians and enable patients to obtain early diagnosis and timely treatment, thereby reducing the mortality rate of pneumonia.

3.3 Algorithm and Process Design

A. Loading the dataset

The dataset is divided into 3 folders: train, test, and val, with subfolders for each visual category (pneumonia and normal).

B. Data visualization and pre-processing

The data seems to be imbalanced therefore to increase the number of training examples we shall use data augmentation.
Matplot is used for previewing the images and dividing it into 20X20 squares. We will perform a grayscale normalization for reducing the effect of illumination's differences. Herewe would be normalizing the data by dividing it by 255 as it’s the maximum pixel of an image.

C. Data augmentation

We need to artificially increase our dataset to prevent the overfitting problem. To recreate the variances, the answer is to change the training data using minor modifications. Horizontal flips, grayscales, random cuts, rotations, and other data augmentation techniques are some of the most common. We can simply increase the number of training instances by using a few of these modification strategies to our training data.

D. Training the model

The experiments were performed on the NVIDIA GeForce GTX 1660 Super configuration. With Cuda acceleration, processing 32 images per batch, the initial learning rate is set to 0. 000001, and the 12 epochs and the classification uses the binary cross-entropy (CE) lossfunction.

3.4 Data Flow Diagrams

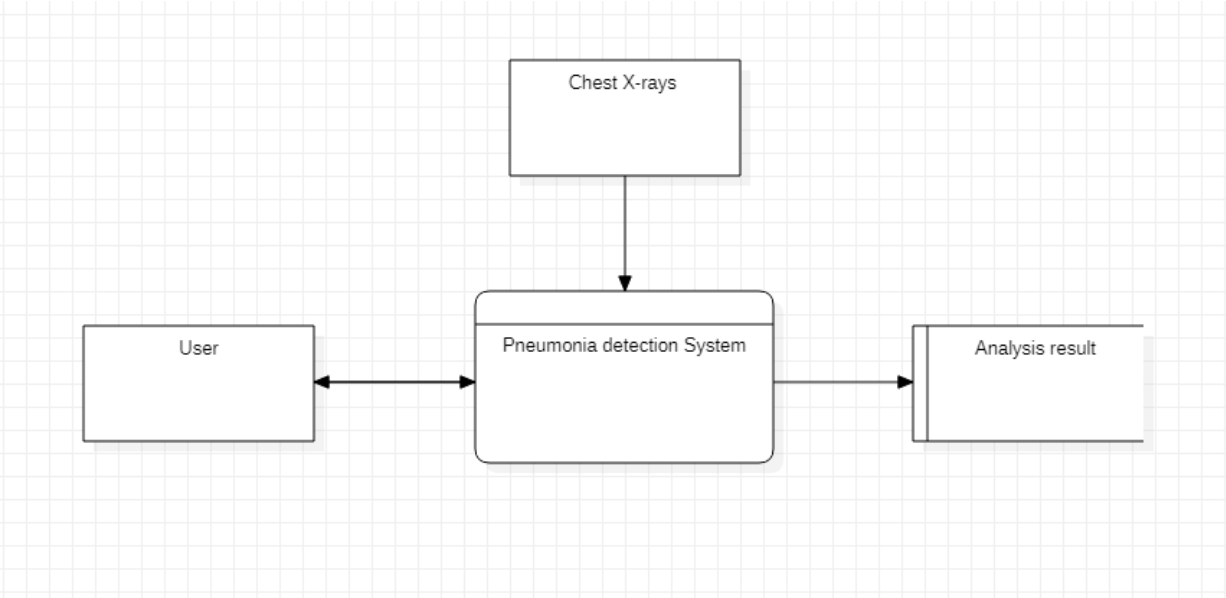


Fig 3.4 Data Flow Diagram (Level 0)

CHAPTER 4
ANALYSIS

ANALYSIS

4.1 Methodology and Implementation

4.1.1 Methodology

In our model, we'll use the data from x-ray pictures. Parts of the lungs affected by pneumonia would be recognised using computer vision. This research proposes an optimal approach for detecting pneumonia using chest X-rays. Our approach is based on a chest X-ray picture collection. The proposed pneumonia detection system uses the Custom Sequential Model. The aforementioned model's design has been broken down into three traditional stages: preprocessing, augmentation, and classification.

4.1.2 Implementation

A. Dataset

Chest Radiograph x-ray (CXR) pictures were collected from the Kaggle website for this research. The dataset is categorized into four major folders: training, testing, and validation, with two subfolders for pneumonia (P) and normal (N) chest X-ray pictures in each of them. There are a total of 5,856 X-ray images of chests, with 4,273 Pneumonia photos and 1,583 Normal images.

B. Convolutional Neural Network

Convolutional neural networks (CNNs, or ConvNets) are a type of deep neural network used to analyze visual images. CNN is a form of deep neural network (DNN) that specializes in image processing and achieves better illness diagnosis accuracy than previous techniques.

It's frequently used in computer vision applications including clustering, object recognition, and picture classification as a result. It's a Deep Learning system that can take an image as an input, assign significance to distinct parts of the picture, and distinguish between them. CNN models include AlexNet, VGG-Net, GoogleNet, and ResNet. The number of convolution layers implemented in each of these models is varied. The more convolution layers in a CNN model, the better the classification accuracy. CNN employs a minimal amount of pre-processing in comparison to other image classification methods. This implies the network will learn the filters that were previously hand-engineered in traditional methods. The ability to develop features without relying on past knowledge or skilled effort is a significant benefit.

C. Preprocessing and Augmentation

In most picture classification applications, the major aim of utilizing a Convolutional Neural Network is to minimize the computational complexity of the model, which is likely to rise if the input is taken as images. To decrease expensive calculation and speed up processing, the original 3-channel pictures will be reduced from 1024 × 1024 to 224 × 224 pixels.

4.2 Feasibility Study

Technical feasibility: Technical feasibility focuses more on the technical resources (software and hardware) available to any organization/companies and also helps to determine whether the technical team of any organization/companies is capable of converting the ideas/methodology into working systems. The software required for our project are Jupyter notebook/google colab and visual studio which are easily available. Hardware requirement is laptop with better processing power which can be managed easily.

Economic feasibility: Economic feasibility assessment typically involves a cost/ benefits analysis of the project. This project developed in minimal amounts will emerge as a huge market valued project. the project focuses on automation of the recruitment process which will help in saving the time and resources of the companies.

Legal feasibility: Legal Feasibility assessment means investigating that whether any aspect of the proposed methodology of a project conflicts with legal requirements like zoning laws, data theft law, etc. The project does not involve any legal concerns since only the required and known data of the user we are storing in a database and making a profile of it on the portal.

Operational feasibility: Operational Feasibility assessment involves undertaking a complete study to understand and analyze how well any organization's needs can be met by completing and implementing the project. The main objective of the project is tomake the recruitment process fast and easy. Automating the process will definitely savethe time of the business sector which is precious.

4.3 Gantt Chart

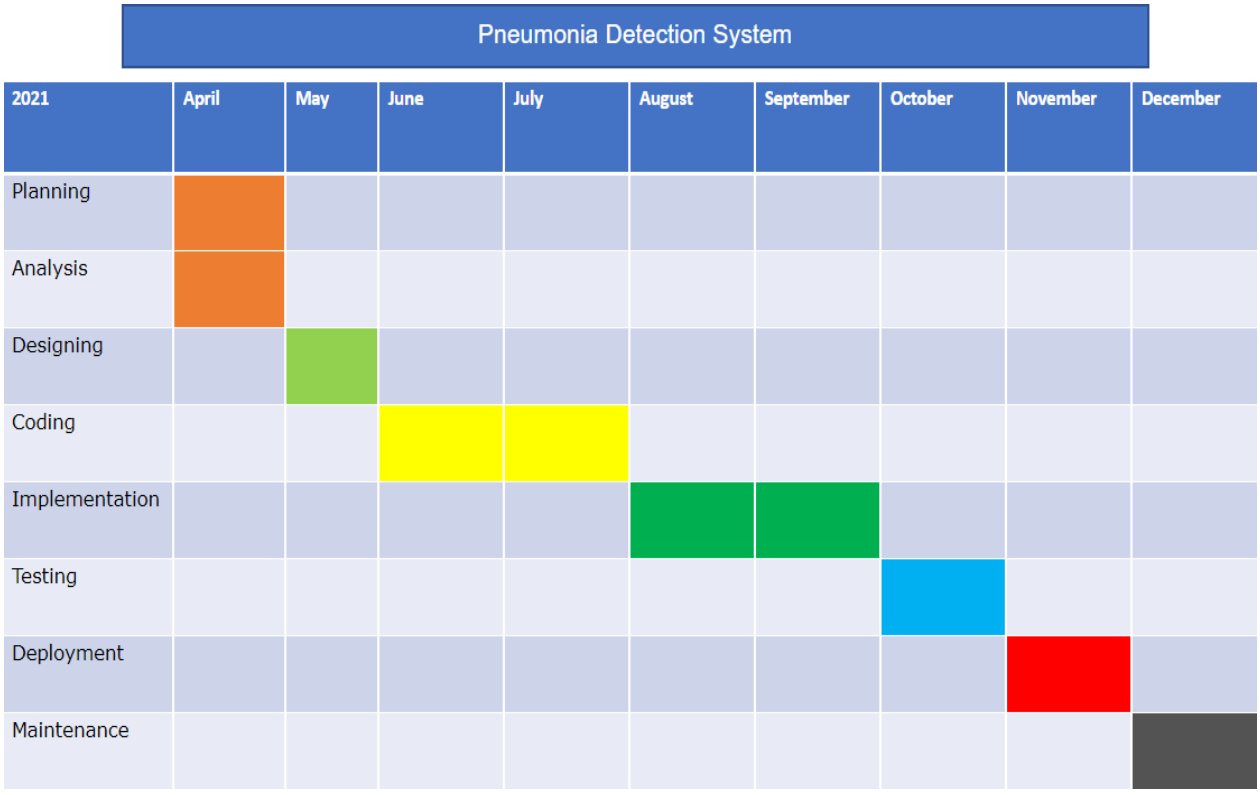


Fig 4.3 Gantt Chart

CHAPTER 5
DESIGN

DESIGN

5.1 UML Diagrams

5.1.1 Use Case Diagram

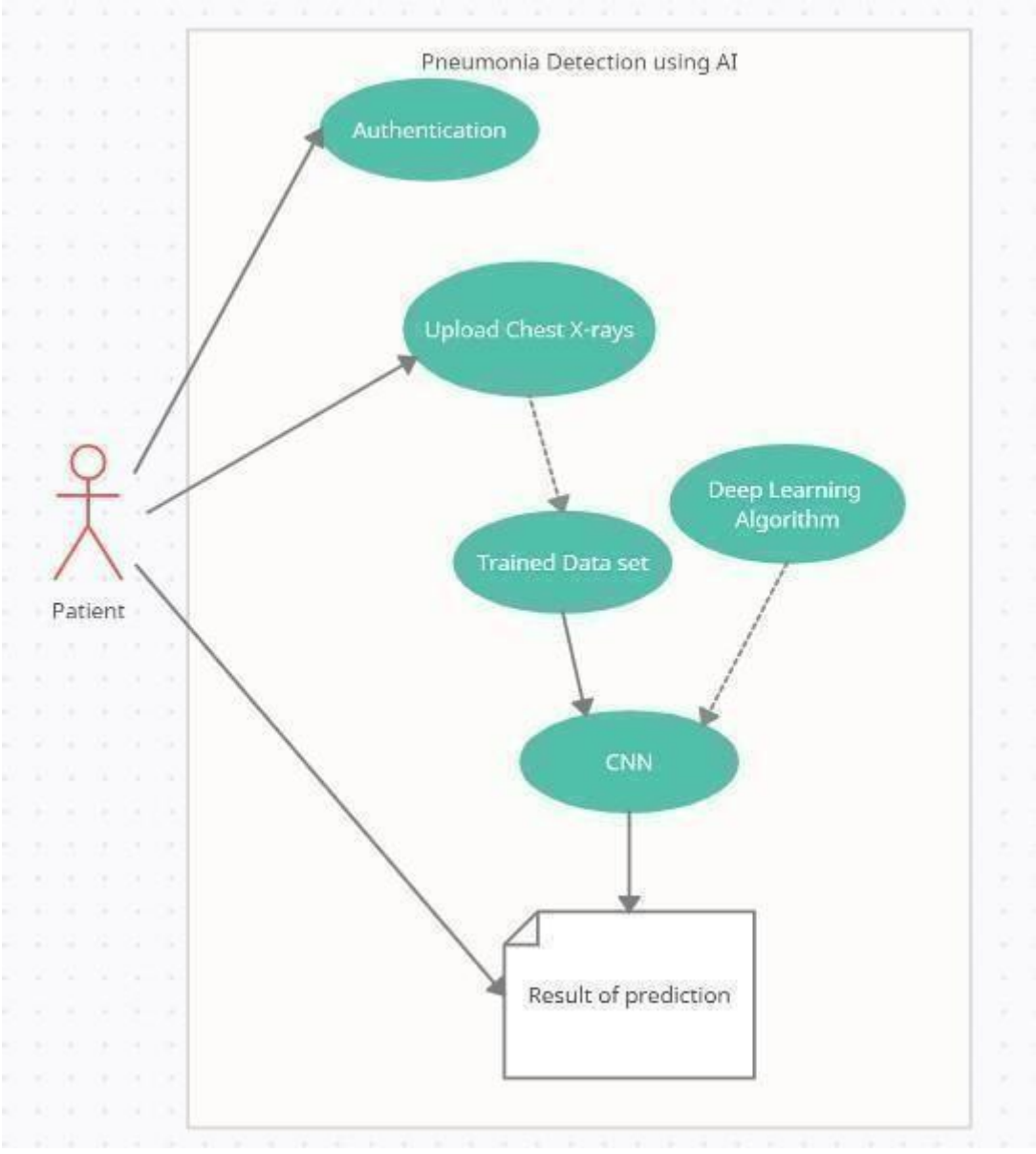


Fig 5.1.1 Use Case Diagram

5.1.2 Activity Diagram

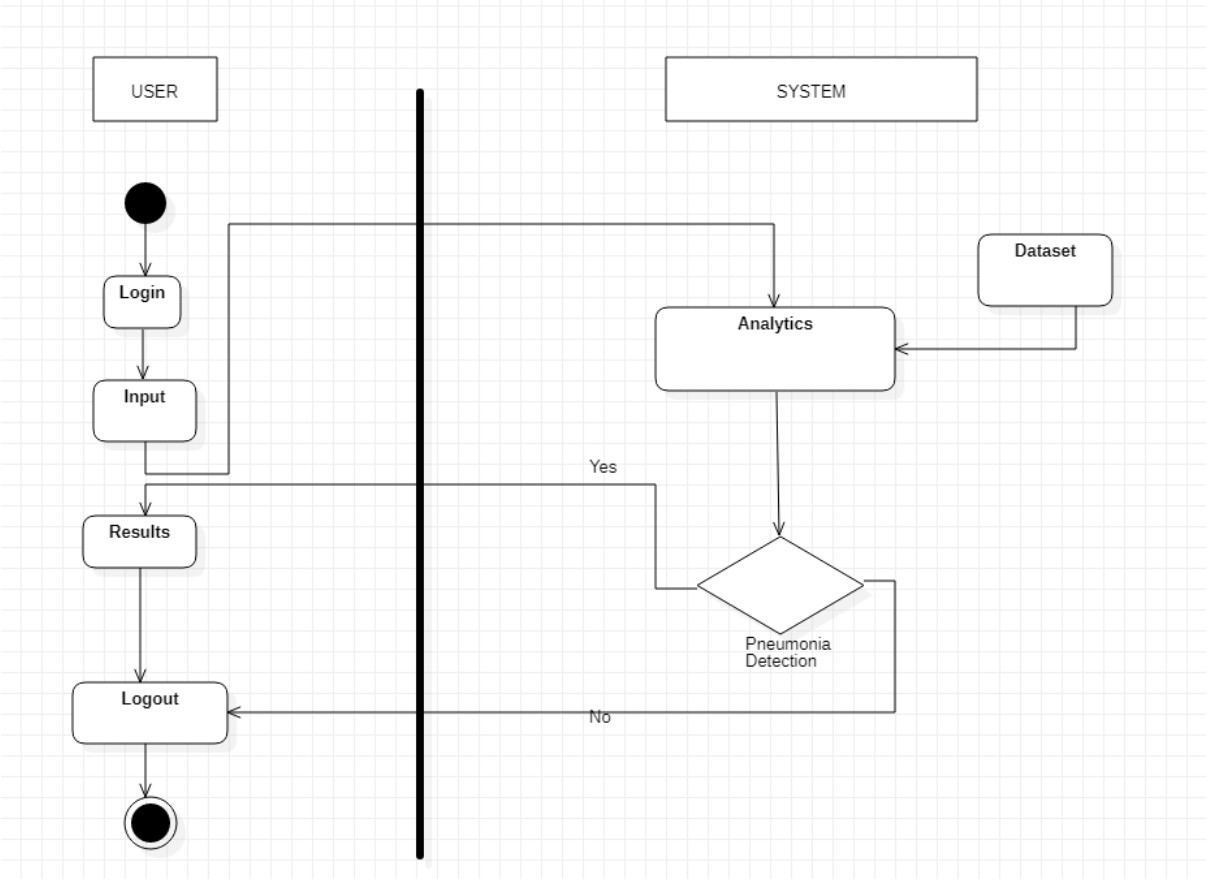


Fig 5.1.2 Activity Diagram

5.1.3 Sequence Diagram

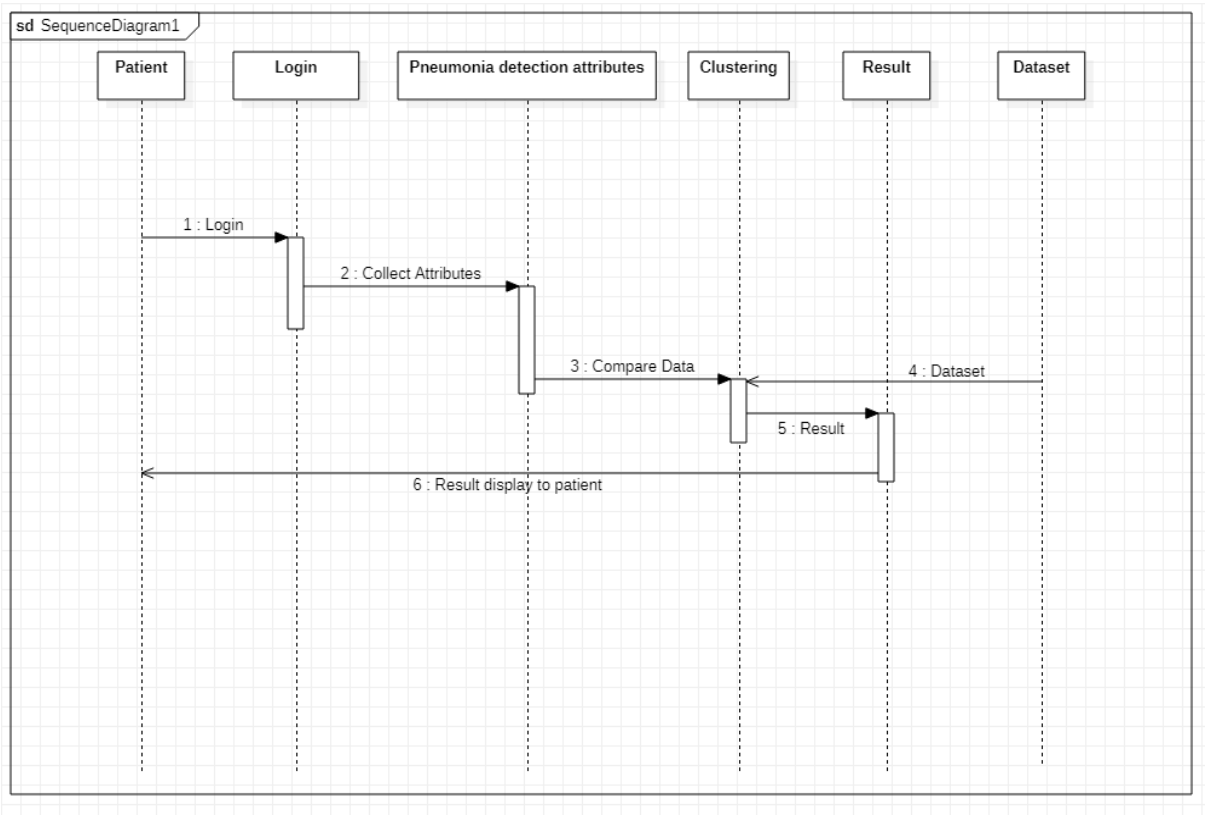


Fig 5.1.3 Sequence Diagram

5.2 Hardware and Software Requirements

5.2.1 Software Requirements:

- Jupyter Notebook
- Code Editor(Visual Studio, Atom)

5.2.2 Hardware Requirements:

- Processor: Intel Pentium 4 or more
- Ram: 1 GB or more
- Hard disk: 40 GB hard disk recommended for the primary partition

CHAPTER 6

RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

6.1 Home page

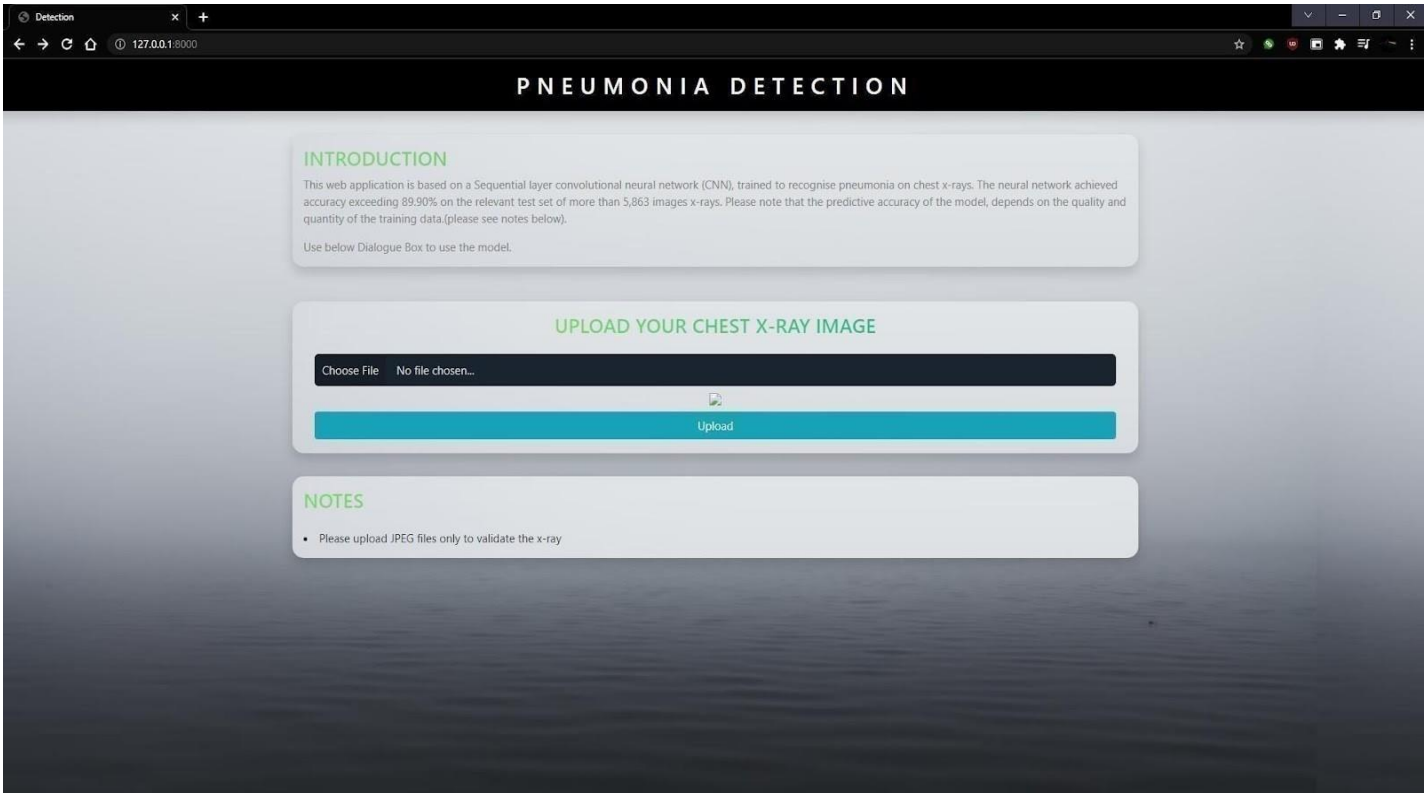


Fig 6.1 Home page

6.2 Uploading the image file

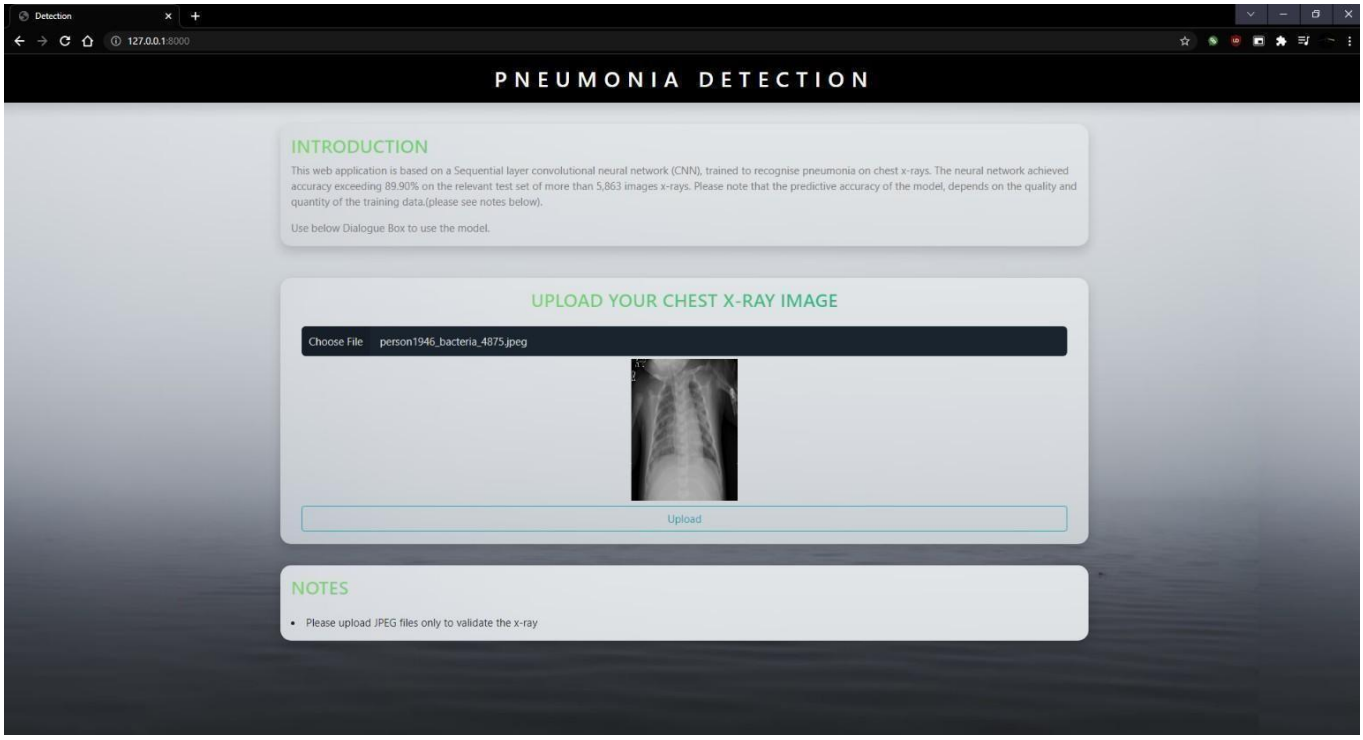


Fig 6.2 Uploaded Image file

6.3 Final Output

A. When a Normal X-ray is uploaded from the validation folder the output given by the model is Normal.

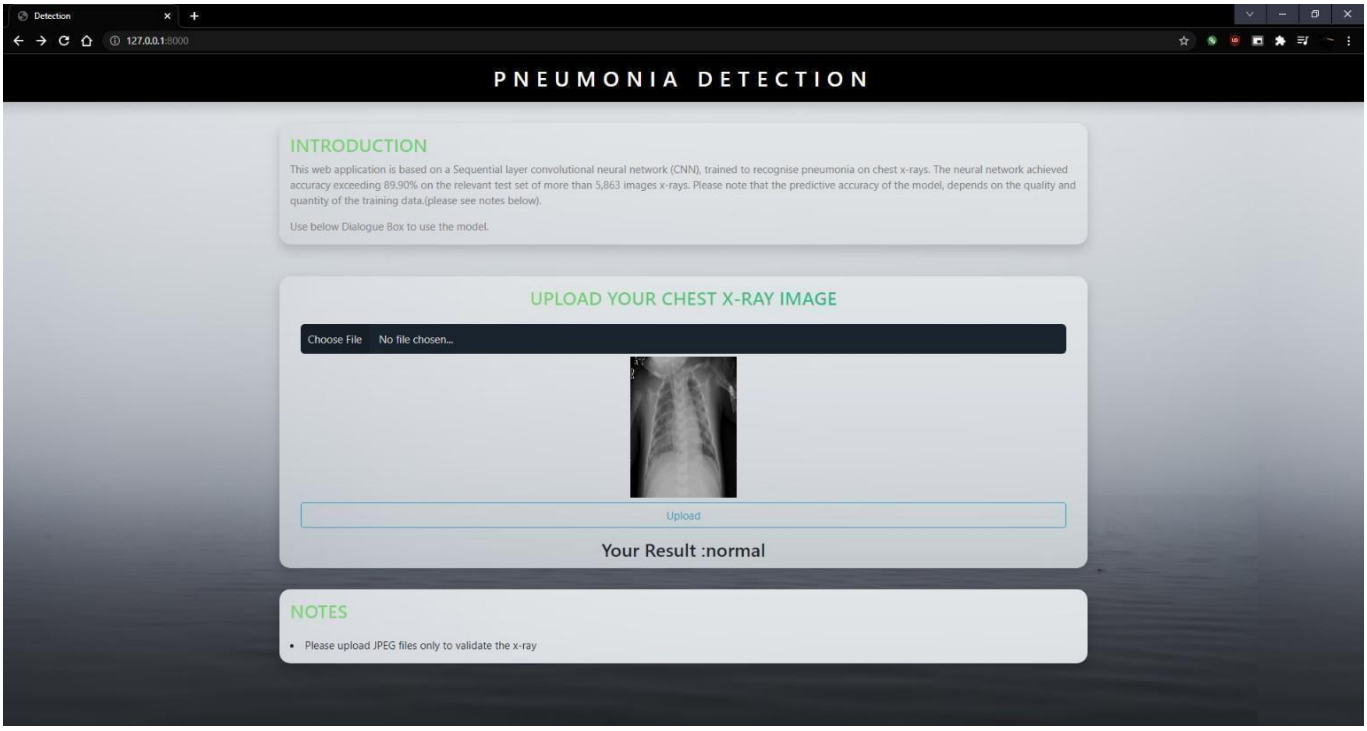


Fig 6.3 Final Output(Normal)

B. When a Pneumonia X-ray is uploaded from the validation folder the output given by the model is Pneumonia detected.

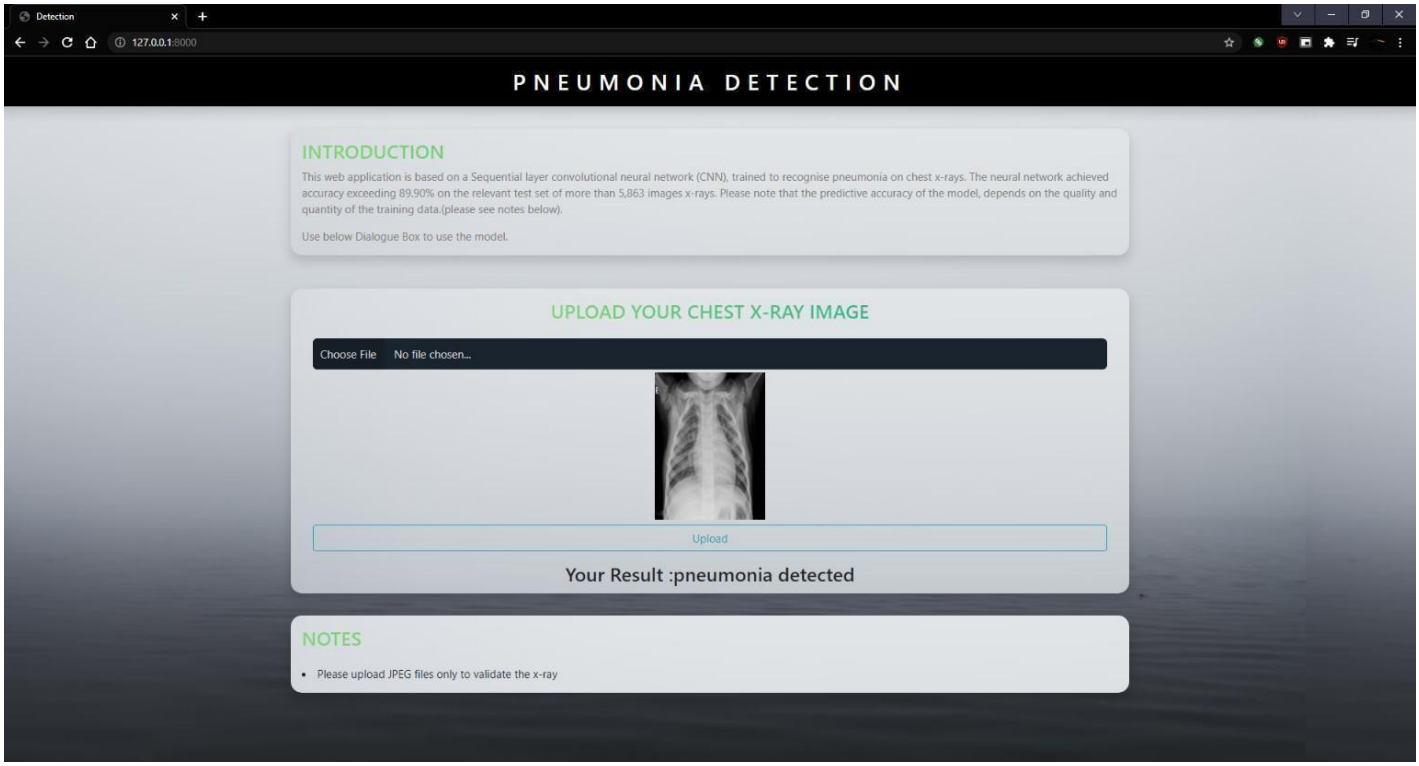


Fig 6.4 Final Output(Pneumonia)

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

To a certain extent, machines can emulate humans. As a result, even the most sophisticated computers and technology cannot always completely replace the function of experts in a certain sector. Although the presence of itinerant medical professionals is required, the presence of experienced radiologists is unavoidable thanks to the advancement of AI technology. Early diagnosis of pneumonia can save a lot of lives and lessen the strain on our healthcare system.

A model has been presented that identifies all positive and negative pneumonia data from a set of X-ray pictures. This technology will assist medical staff in making real-time decisions.

7.2 Future Scope

For improvements in the project we can adapt to other convolutional neural network architectures like Inception-v3, shuffle Net, and Mobile Net architectures for pneumonia detection can be implemented and the optimization of hyper-parameters should also be considered to improve the accuracy of the model.Also, the performance of the proposed Deep CNN model can be improved with more number of layers and parameters.

This project can help medical workforce in their decision making for a real-time application of the use of accurate models in detecting pneumonia and discover the potential of diagnosing pneumonia using deep learning. Different lung disease classes can also be included in this model to detect various lung diseases using the chest X-ray images.This will allow clinicians to recognize lung diseases from chest X-ray images with lower prevalence at an earlier stage of the disease.

REFERENCE

REFERENCE

[1] WHO, “World pneumonia day”, 2018.

[2] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, and Ronald M Summers, “Chest x-ray 8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of the common lung diseases,” 2017.

[3] Ahmed Tahseen Minhaz, and Khalid Ashraf, “Abnormality detection and localization in chest x-rays using deep convolutional neural networks,” 2017 .

[4] A. Bilbily, M. Cicero et al., “Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs,” 2017

[5] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, “Survey on deep learning for pulmonary medical imaging,” *Front Medicine*, vol. 14, no. 4, 2019.

[6] P. Tiwari, D. Gupta, A. K. Jaiswal, S. Kumar, A. Khanna, “Identifying pneumonia in chest X-rays: a deep learning approach,” *Measurement*, vol. 145, pp. 511–518, 2019.

[7] Brandon Yang, Jeremy Irvin, Pranav Rajpurkar, Hershel Mehta, Daisy Ding, Katie Shpanskaya, and others, “CheXnet: Radiologist-level pneumonia detection on chest x rays with Deep Learning,” 2017.

[8] Joshua Kravitz, Benjamin Antin, and Emil Martayan,

“Detecting Pneumonia in Chest X-Rays with Supervised Learning,” 2017.

[9] Okele Stephen, ”An Efficient Deep Learning

Approach to Pneumonia Classification in Healthcare,” 2019

PAPER PUBLISHED

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 09 Issue: 04 | Apr 2022 www.irjet.net p-ISSN: 2395-0072

Pneumonia Detection System using AI

[1]Neha Kudu, [2]Manasi Sherkhane, [3]Omkar Wadekar, [4]Ritesh Rahatal

[1,2,3,4]Department of Information Technology, Vidyalankar Institute of Technology, Mumbai, India

---

***Abstract*—** Pneumonia is known as one of the most life-threatening diseases in the world which affects the lung(s) of humans and is known as one of the leading causes of death in India. Roughly 33% of the deaths in India are caused as a result of pneumonia and as reported by the World Health Organization (WHO). Currently pneumonia is diagnosed using a Chest X-Ray image which is then evaluated by an expert radiotherapist. This process is quiet exerting and travail and it often leads to a difference in opinion among the experts. Thus developing a solitary system would be beneficial for identification, preventing further transmission and treatment in remote areas. This system proposes a cnn model which has been trained from scratch and that will classify and also detect the presence of pneumonia from a dataset of chest X-ray images. For this system the cnn model would extract features from a given dataset of chest X-ray image and then classify it to work it out if an individual is infected with pneumonia or not.

Keywords*—Pneumonia, Chest X-ray images, Deep learning, convolutional Neural Network.*

I. INTRODUCTION

Pneumonia is a disease that disturbs the alveoli of the lungs and has a mortal account which accounts to about 16% of the world deaths [1], being the world's leading cause for deaths among many children. Pneumonia is responsible for almost 127,000 deaths in India the numbers are rising due to the spread of novel coronavirus. The risk of pneumonia is quite high, especially in developing countries where billions of people live in energy poverty and rely on polluting energy sources. Each year, nearly 4 million people die prematurely as a result of diseases related by indoor air pollution, suchas pneumonia, according to the World Health Organization.On a yearly basis, about 150 million individuals, mostly children under the age of five, become sick with pneumonia. Chest X-ray (CXR) is the most suitable imaging modality to diagnose pneumonia among others. Pneumonia manifests as an area or areas of increased opacity on CXR. In particular, convolutional neural networks (CNN)-based approaches have been effectively used to categorise illnesses and detect aberrant areas or segment lesions in CXR pictures. A chest x-ray is one of the most common medical methods used to identify the illness.. As a focused beam of electrons known as x- ray photons passes through the human tissues, a picture is created on a metal surface known as photographic film. When a radiologist examines chest X-rays for Pneumonia, he or she will look for white spots in the lungs called infiltrates, which suggest an infection. The restricted colour palette of x-ray pictures consists of just black and white colours, which has limitations for detecting whether an infected region in the lungs exists.

II. LITERATURE SURVEY

The latest improvements in the field of Machine Learning and AI mainly due to large scale usage of Convolutional Neural Networks (CNNs) and the availability of free dataset. That was once considered to be very rare and has assisted various algorithms to perform much better that was not considered to be a commonplace a few years ago. The automated diagnosis of varied diseases has received a high interest. The low performance of several CNN models on diverse abnormalities proves that a single model cannot be used for all the purposes. So for a better exploration of machine learning in the chest screening, Wang et al. (2017) [2] released a larger dataset of frontal chest X-Rays.

Huang et al. [3] adopted deep learning techniques. Performance of different variants of Convolutional Neural Networks for abnormality detection in chest X- Rays was then proposed by Islam et al. using the publicly available OpenI dataset for the better exploration of machine learning in chest screening.

Cicero et al. discussed the training and validation of CNNs with modest-sized medical data to detect pathology in 2017 [4]. Ma et al. presented a survey on deep learning for pulmonary medical imaging in 2019 [5]. Jaiswal et al. from University of Bedfordshire described an approach based on deep learning for identifying pneumonia in chest X-way in 2019 [6].In recent times, Pranav Rajpurkar, Jeremy Irvin, et al. (2017) [7] exploredthe dataset fordetecting pneumonia at a level far better than radiologists, they referred their model as ChexNet that uses DenseNet- 121 layer architecture for detecting all the 14 diseases from a lot of 112,200 images that are available in the dataset. After theCheXNet [7] model, Benjamin Antin et al. (2017) [8] worked on the same dataset and proposed a logistic regression model for detecting pneumonia.

Okeke Stephen, Mangal Sain , Uchenna, and Do-Un Jeong In [9] proposed a cnn model **t**hat's trained from scratch to classify and to detect the presence of pneumonia from a collection of chest X-ray image dataset. Unlike other methods that rely only on transfer learning approaches **to** realize an interesting classification performance, they constructed a convolutional neural network model from scratch that extract features from a given chest X-ray image and classify it to work out if an individual is infected with pneumonia. This model could help mitigate the reliability and interpretability challenges often faced when handling medical imagery.

III. PROPOSED SYSTEM

This method presents a convolutional neural network model that can identify and detect the presence of pneumonia from a set of chest X-ray image samples after being trained from scratch. In our model, we'll use the data from x-ray pictures. We'll build a convolutional neural network model from the ground up to extract characteristics from a chest X-ray picture and then categorize it to see if a person has pneumonia or not. Parts of the lungs afflicted by pneumonia would be recognized using computer vision.This research proposes an optimal approach for detecting pneumonia using chest X-rays.

IV. METHODOLOGY

In our model, we'll use the data from x-ray pictures. Parts of the lungs afflicted by pneumonia would be recognised using computer vision. This research proposes an optimal approach for detecting pneumonia using chest X-rays. Our approach is based on a chest X-ray picture collection. The proposed pneumonia detection system uses the Custom Sequential Model. The aforementioned model's design has been broken down into three traditional stages: preprocessing, augmentation, and classification.

A. DATASET

Chest Radiograph x-ray (CXR) pictures were collected from the Kaggle website for this research. The dataset is categorized into four major folders: training, testing, and validation, with two subfolders for pneumonia (P) and normal (N) chest X-ray pictures in each of them. There are a total of 5,856 X-ray images of chests, with 4,273 Pneumonia photos and 1,583 Normal images.

B. CONVOLUTIONAL NEURAL NETWORK

Convolutional neural networks (CNNs, or ConvNets) are a type of deep neural network used to analyse visual images. CNN is a form of deep neural network (DNN) that specialises in image processing and achieves better illness diagnosis accuracy than previous techniques.

It's frequently utilised in computer vision applications including clustering, object recognition, and picture classification as a result. It's a Deep Learning system that can take an image as an input, assign significance to distinct parts of the picture, and distinguish between them. CNN models include AlexNet, VGG-Net, GoogleNet, and ResNet. The number of convolution layers implemented in each of these models is varied. The more convolution layers in a CNN model, the better the classification accuracy. CNN employs a minimal amount of pre-processing in comparison to other image classification methods. This implies the network will learn the filters that were previously hand-engineered in traditional methods. The ability to develop features without relying on past knowledge or skilled effort is a significant benefit.

C. PREPROCESSING AND AUGMENTATION

In most picture classification applications, the major aim of utilizing a Convolutional Neural Network is to minimize the computational complexity of the model, which is likely to rise if the input is taken as images. To decrease expensive calculation and speed up processing, the original 3-channel pictures will be reduced from 1024 × 1024 to 224 × 224 pixels.

V. ALGORITHM

Loading the dataset.

The dataset is divided into 3 folders: train, test, and val, with subfolders for each visual category (pneumonia and normal).

Data visualization and pre-processing.

The data seems to be imbalanced therefore to increase the number of training examples we shall use data augmentation.

Matplot is used for previewing the images and dividing it into 20X20 squares. We will perform a grayscale normalization for reducing the effect of illumination's differences. Here we would be normalizing the data by dividing it by 255 as it's the maximum pixel of an image.

Data augmentation.

We need to artificially increase our dataset to prevent the overfitting problem. To recreate the variances, the answer is to change the training data using minor modifications. Horizontal flips, grayscales, random cuts, rotations, and other data augmentation techniques are some of the most common. We can simply increase the number of training instances by using a few of these modification strategies to our training data.

Training the model

The experiments were performed on the NVIDIA GeForce GTX 1660 Super configuration. With Cuda acceleration, processing 32 images per batch, the initial learning rate is set to 0. 000001, and the 12 epochs and the classification uses the binary cross- entropy (CE) loss function.

Analysis after model training

Precision is the ratio between the True Positives and all the Positives. For our problem statement, that would be the measure of patients that we correctly identify having a Pneumonia out of all the patients actually having it. Mathematically:

$$\text{Precision} - \frac{True\ Positive\ (TP)}{True\ Positive\ (TP)+False\ Positive(FP)}$$

Recall is the measure of our model correctly identifying True Positives. Thus, for all the patients who actually have heart disease, recall tells us how many we correctly identified as having a heart disease. Mathematically:

$$\text{Recall} - \frac{True\ Positive\ (TP)}{True\ Positive\ (TP)+False\ Negative(FN)}$$

Accuracy is the ratio of the total number of correct predictions and the total number of predictions

$$\text{Accuracy-} \frac{True\ Positive\ (TP)+True\ Negative(TN)}{True\ Positive + False\ Positive + True\ Negative(TN)+False\ Negative}$$

Role of the F1-Score

F1-score is the Harmonic mean of the Precision and Recall:

$$\text{F1-Score} - 2 * \frac{Precision * Recall}{Precision + Recall}$$

VI. CONCLUSION

To a certain extent, machines can emulate humans. As a result, even the most sophisticated computers and technology cannot always completely replace the function of experts in a certain sector. Although the presence of itinerant medical professionals is required, the presence of experienced radiologists is unavoidable thanks to the advancement of AI technology. Early diagnosis of pneumonia can save a lot of lives and lessen the strain on our healthcare system.A model has been presented that identifies all positive and negative pneumonia data from a set of X-ray pictures. This technology will assist medical staff in making real-time decisions.

REFERENCES

[1] WHO, "World pneumonia day", 2018.

[2] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, and Ronald M Summers, "Chest x-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of the common lung diseases," 2017.

[3] Ahmed Tahseen Minhaz, and Khalid Ashraf, "Abnormality detection and localization in chest x-rays using deep convolutional neural networks," 2017 .

[4] A. Bilbily, M. Cicero et al., "Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs," 2017

[5] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, "Survey on deep learning for pulmonary medical imaging," *Front Medicine*, vol. 14, no. 4, 2019.

[6] P. Tiwari, D. Gupta, A. K. Jaiswal, S. Kumar, A. Khanna, "Identifying pneumonia in chest X-rays: a deep learning approach," *Measurement*, vol. 145, pp. 511–518, 2019.

[7] Brandon Yang, Jeremy Irvin, Pranav Rajpurkar, Hershel Mehta, Daisy Ding, Katie Shpanskaya, and others, "CheXnet: Radiologist-level pneumonia detection on chest x rays with Deep Learning," 2017.

[8] Joshua Kravitz, Benjamin Antin, and Emil Martayan, "Detecting Pneumonia in Chest X-Rays with Supervised Learning," 2017.

[9] Okele Stephen, "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare," 2019

CERTIFICATIONS


e-ISSN: 2395-0056 p-ISSN: 2395-0072
International Research Journal of Engineering and Technology (IRJET)
(An ISO 9001 : 2008 Certified Journal)
Is hereby awarding this certificate to
Neha Kudu
In recognition the publication of the manuscript entitled
Pneumonia Detection System using AI
published in our Journal Volume 9 Issue 4 April 2022
Editor in Chief
E-mail : editor@irjet.net
Impact Factor : 7.529
www.irjet.net


e-ISSN: 2395-0056 p-ISSN: 2395-0072
International Research Journal of Engineering and Technology (IRJET)
(An ISO 9001 : 2008 Certified Journal)
Is hereby awarding this certificate to
Omkar Wadekar
In recognition the publication of the manuscript entitled
Pneumonia Detection System using AI
published in our Journal Volume 9 Issue 4 April 2022
Editor in Chief
E-mail : editor@irjet.net
Impact Factor : 7.529
www.irjet.net


e-ISSN: 2395-0056 p-ISSN: 2395-0072
International Research Journal of Engineering and Technology (IRJET)
(An ISO 9001 : 2008 Certified Journal)
Is hereby awarding this certificate to
Ritesh Rahatal
In recognition the publication of the manuscript entitled
Pneumonia Detection System using AI
published in our Journal Volume 9 Issue 4 April 2022
Editor in Chief
E-mail : editor@irjet.net
Impact Factor : 7.529
www.irjet.net


e-ISSN: 2395-0056 p-ISSN: 2395-0072
International Research Journal of Engineering and Technology (IRJET)
(An ISO 9001 : 2008 Certified Journal)
Is hereby awarding this certificate to
Manasi Sherkhane
In recognition the publication of the manuscript entitled
Pneumonia Detection System using AI
published in our Journal Volume 9 Issue 4 April 2022
Editor in Chief
E-mail : editor@irjet.net
Impact Factor : 7.529
www.irjet.net