DETECTION AND COMPARISON OF ALGORITHM FOR PREDICTION OF PNEUMONIA

A PROJECT REPORT

Submitted by

RAGUL R (312419104103)

RIFFA HAMMED (312419104110)

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

St. JOSEPH'S INSTITUTE OF TECHNOLOGY

(An Autonomous Institution)

ANNA UNIVERSITY :: CHENNAI 600025

MARCH 2023

ANNA UNIVERSITY : CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report "**DETECTION AND COMPARISON OF ALGORITHM FOR PREDICTION OF PNEUMONIA** " is the bonafide work of "**RAGUL R(312419104103)** and **RIFFA HAMMED (312419104110)"** who carried out the project under my supervision.

SIGNATURE	**SIGNATURE**
Dr. J. DAFNI ROSE, M.E, Ph.D.,	**Mrs. M.V.EZHIL DYANA, B.E., M.E.,**
PROFESSOR AND HEAD,	**SUPERVISOR,**
	Assistant Professor,
Computer Science and Engineering,	**Computer Science and Engineering,**
St. Joseph's Institute of Technology,	St. Joseph's Institute of Technology,
Old Mamallapuram Road,	Old Mamallapuram Road,
Chennai - 600119.	Chennai - 600119.

ACKNOWLEDGEMENT

We also take this opportunity to thank our respected and honourable Chairman **Dr. B. Babu Manoharan M.A., M.B.A., Ph.D.,** for the guidance he offered during our tenure in this institution.

We extend our heartfelt gratitude to our respected and honourable Managing Director **Mrs. S. Jessie Priya M.Com.,** for providing us with the required resources to carry out thisproject.

We express our deep gratitude to our honourable Executive Director **Mr. B. Sashi Sekar M.Sc. (INTL. Business)** for the constant guidance and support for ourproject.

We are indebted to our Principal **Dr. P. Ravichandran M.Tech., Ph.D.,** for granting us permission to undertake this project.

We would like to express our earnest gratitude to our Head of the Department **Dr. J. Dafni Rose M.E., Ph.D.,** for her commendable support and encouragement for the completion of the project with perfection.

We also take the opportunity to express our profound gratitude to our guide **Mrs.M.V.Ezhil Dyana B.E, M.E,** for her guidance, constant encouragement, immense help and valuable advice for the completion of this project.

We wish to convey our sincere thanks to all the teaching and non- teaching staff of the department of **COMPUTER SCIENCE AND ENGINEERING** without whose cooperation this venture would not have been a success.

CERTIFICATE OF EVALUATION

College Name : St. JOSEPH’S INSTITUTE OF TECHNOLOGY
Branch : COMPUTER SCIENCE AND ENGINEERING
Semester :VIII

Sl.No	Name of the Students	Title of the Project	Name of the Supervisor with designation
1 2	RAGUL R (312419104103) RIFFA HAMMED (312419104110)	DETECTION AND COMPARISON OF ALGORITHM FOR PREDICTION OF PNEUMONIA	MRS.EZHIL DYANA, B.E, M.E., Assistant Professor

The report of the project work submitted by the above students in partial fulfilment for the award of Bachelor of Engineering Degree in Computer Science and Engineering of Anna University were evaluated and confirmed to be report of the work done by above students.

Submitted for project review and viva voce exam held on ______________

(INTERNAL EXAMINER) **(EXTERNAL EXAMINER)**

ABSTRACT

Pneumonia has caused significant deaths worldwide, and it is a challenging task to detect many lung diseases such as like atelectasis, cardiomegaly, lung cancer, etc., often due to limited professional radiologists in hospital settings. This lung disease is more common in people older than 65 and children under five years old. Although the treatment of pneumonia can be challenging, it can be prevented by early diagnosis using Computer-Aided Diagnosis (CAD) systems. Chest X-Rays (CXRs) are currently the primary imaging tool for detection of pneumonia, which are widely used by radiologists. In this regard, a novel hybrid Convolutional Neural Network (CNN) model is proposed using three classification approaches. In the first classification approach, Fully-Connected (FC) layers are utilized for the classification of CXR images. The pre-processing required in a ConvNet is much lower as compared to other classification compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/ characteristics. This model is trained for several epochs and the weights that result in the highest classification accuracy are saved. In the second classification approach, the trained optimized weights are utilized to extract the most representative CXR image features and Machine Learning (ML) classifiers are employed to classify the images. In the third classification approach, an ensemble of the proposed classifiers is created to classify CXR images. The results suggest that the proposed ensemble classifier using Support Vector Machine (SVM) with Radial Basis Function (RBF) and Logistic Regression (LR) classifiers has the best performance with 98.55% accuracy.

TABLE OF CONTENT

FIGURE.NO	LIST OF FIGURES	PAGE NO
3.1	Use Case Diagram for Pneumonia Prediction Using Algorithms.	14
3.2	Sequence Diagram for Pneumonia Prediction Using Algorithms.	15
3.3	Activity Diagram for Pneumonia Prediction Using Algorithms.	16
3.4	Collaboration Diagram for Pneumonia Prediction Using Algorithms.	17
4.1	Architecture Diagram for Pneumonia Prediction Using Algorithms.	18
6.2.1	Jupyter Notebook	34
6.2.2	Pneumonia Dataset	34
6.2.3	ROC	35
6.2.4	Classification Report	39
6.2.5	Comparison Of Algorithms	43
6.2.6	Confusion Matrix	44
6.2.7	Accuracy Of Models	46
6.3	PERFORMANCE ANALYSIS	50

LIST OF ABBREVIATIONS

ACRONYM	ABBREVIATION
CXR	Chest X-Ray
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
SVM	Support Vector Machine
CNN	Convolutional Neural Network
DL	Deep Learning
CAD	Computer-Aided Diagnosis
FC	Fully Connected
SVM	Support Vector Machine
RBF	Radial Basis Function
LR	Logistic Regression
UML	Unified Modeling Language
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
VGG	Visual Geometry Group
RESNET	Residual Network
DHE	Dynamic Histogram Equalization
HE	Histogram Equalization
EHR	Electronic Health Record

CHAPTER 1
INTRODUCTION

Pneumonia is a lung disease that is caused by acute respiratory infection. Pneumonia causes reduced oxygen intake and painful breathing. Although pneumonia can affect people at any age, it is more common in people older than 65 and children under five years old. Some popular tools to evaluate the presence of pneumonia in a person are Chest X-Ray (CXR), Computed Tomography (CT) of the lungs, ultrasound of the chest, Magnetic Resonance Imaging (MRI) of the chest, and needle biopsy of the lung. Medical X-rays are electromagnetic radiations that have higher energy than visible light and can penetrate through most objects.

1.1 OVERVIEW

Lung Cancer may be a noteworthy reason for Mortality within the western world as exhibited by the striking factual numbers distributed consistently by the American Carcinoma Society. They demonstrate that the 5-year survival rate for patients with lung malignancy are often enhanced from a standard of 14% up to 49% if the ailment is analyzed and treated at its initial stage. Medicinal pictures as a significant piece of therapeutic determination and treatment were specializing in these pictures permanently. These pictures incorporate success of concealed data that misused by doctors in selecting contemplated choices around a patient. Then again, removing this important shrouded data may be a basic first stride to their utilization. This reason inspires to utilize information digging systems abilities for productive learning extraction & find concealed lung. Mining Medical Pictures includes numerous procedures. Medicinal data processing may be a promising zone of computational insight connected to a consequently break down patient's records going for the disclosure of latest information valuable for restorative choice making. Affected information is anticipated not just to increment exact determination and effective infection treatment, additionally to enhance security by diminishing blunders. The systems during this paper arrange the advanced X-beam midsection movies in two classes: ordinary and strange. The normal ones are those portraying a solid patient. The irregular ones incorporate style of lung tumor; we'll utilize a typical arrangement technique specifically SVMs & neural systems.

1.2 PROBLEM STATEMENT

Obtaining high-quality and diverse chest X-ray images for training and testing the CNN model can be a challenge. It is important to ensure that the dataset used for training and testing is large and representative of the population for which the model is intended. In addition, it is important to ensure that the data is of high quality, with consistent imaging techniques and appropriate annotations. Any use of medical data for research purposes must comply with ethical guidelines and obtain appropriate approvals. It is important to ensure that the use of chest X-ray images for pneumonia detection is conducted with informed consent, and that the privacy and confidentiality of the patients is protected.

1.3 EXISTING SYSTEM

Although the treatment of pneumonia can be challenging, it can be prevented with modest treatments and cured with low-cost, low-tech medication and care. Therefore, is an urgent need to develop diagnostic tools in order to reduce pneumonia-related mortality, particularly in children and old people. Chest X-Ray (CXR), Computed Tomography (CT) of the lungs, ultrasound of the chest, Magnetic Resonance Imaging (MRI) of the chest, and needle biopsy of the lungs. The past recent years, Convolutional Neural Networks (CNNs) have shown a great potential in image classification and segmentation and are widely utilized for creating DL - based CAD systems. Using a hybrid CNN model for detection of pneumonia is of a significant importance. Moreover, none of the previous studies have implemented a CAD system for a reliable and accurate detection of pneumonia.

1.4 PROPOSED SYSTEM

In proposed system, we have designed the proposed hybrid CNN model with Fully Connected (FC) layers and trained the hybrid model for a defined number of epochs and updated the weights of the model using back propagation process and saved the best-performing weights. To prevent over fitting in training the proposed hybrid CNN model with FC layers, a dropout ratio of 0.5 was deployed after each FC layer during the training process. To train the proposed ensemble classifier, first, the trained hybrid CNN model with FC layers and ML classifiers has to be utilized to extract the features of CXR images in the test set. The highest training accuracy in the proposed hybrid CNN model with FC layers is 97.78%, which was brought about in epoch 20. The results suggest that the proposed ensemble classifier using Support Vector Machine (SVM) with Radial Basis Function (RBF) and Logistic Regression (LR) classifiers has the best performance with 98.55% accuracy. Inception Modules are used in Convolutional Neural Networks to allow for more efficient computation and deeper Networks through a dimensionality reduction with stacked 1×1 convolutions. Xception slightly outperforms Inception v3 on the ImageNet dataset, and vastly outperforms it on a larger image classification dataset with 17,000 classes. Most importantly, it has the same number of model parameters as Inception, implying a greater computational efficiency.ResNets is to build a deeper networks compared to other plain networks and simultaneously find a optimised number of layers to negate the vanishing gradient problem.Even though ResNet is much deeper than VGG16 and VGG19, the model size is actually substantially smaller due to the usage of global average pooling rather than fully-connected layers.

CHAPTER 2
LITERATURE REVIEW

Aniket et .al (2016)[3] Lung cancer mortality rate is the highest among all other types of cancer. It is one of the most serious cancers in the world, with the survival rate very less after the diagnosis. Survival from lung cancer is directly related to its growth at its detection time. The earlier the detection is, the higher the chances of successful treatment are. An estimated 85% of lung Cancer cases in males and 75% in females are caused by cigarette smoking [1]. There are many techniques to diagnose lung cancer ,like Chest Radiography (x-ray), Computed Tomography (CT), Magnetic Resonance Imaging (MRI scan) .but, most of these techniques are costly and time consuming. And most of these techniques are detecting the lung cancer in Its advanced stages. Hence, there is a great need of a new technology to diagnose the lung cancer in its early stages. Image processing techniques provide a good class tool for cultivating the manual analysis[2].

Arnold F.L et.al (2020)[4]Community-acquired pneumonia (CAP) is the eighth leading cause of death in the United States, and in adults aged 65 years and older.1,2 It also causes more death in the United States than any other infectious disease.3 The cumulative cost in the United States for all inpatients with CAP has been estimated to be $9 to $17 billion/year.4-6 The last National Hospital Discharge survey addressing pneumonia as a first-listed International Classification of Diseases, Ninth Revision (ICD-9), diagnosis was for 2007 when 1 056 000 patients were identified, of whom 610 000 were aged 65 years or older.7 The incidence of hospitalization for CAP that is currently known for older adults ranges from 1150 to 1830 per 100 000 population.8-10 Every large epidemiological study has been limited by defining patients with ICD data rather than with clinical and radiographic data.

There are no studies that have included comprehensive CAP data for one city from which values could be translated to the US population as a whole. A comprehensive study is needed. Advancement in this field will come with accurate and timely familiarity of the epidemiology, incidence, and mortality of CAP in older adults.

Attaway A.H et.al (2020)[2] Severe covid-19 pneumonia has posed critical challenges for the research and medical communities. Older age, male sex, and comorbidities increase the risk for severe disease. For people hospitalized with covid-19, 15-30% will go on to develop covid-19 associated acute respiratory distress syndrome (CARDS). Autopsy studies of patients who died of severe SARS CoV-2 infection reveal presence of diffuse alveolar damage consistent with ARDS but with a higher thrombus burden in pulmonary capillaries. When used appropriately, high flow nasal cannula (HFNC) may allow CARDS patients to avoid intubation, and does not increase risk for disease transmission. During invasive mechanical ventilation, low tidal volume ventilation and positive end expiratory pressure (PEEP) titration to optimize oxygenation are recommended. Dexamethasone treatment improves mortality for the treatment of severe and critical covid-19, while remdesivir may have modest benefit in time to recovery in patients with severe disease but shows no statistically significant benefit in mortality or other clinical outcomes. Covid-19 survivors, especially patients with ARDS, are at high risk for long term physical and mental impairments, and an interdisciplinary approach is essential for critical illness recovery.

Chan.H et.al (2020)[5] Medical imaging is an important diagnostic tool for various diseases. Roentgen discovered that X-rays could non-invasively look into the human body in 1895 and X-ray radiography became the first diagnostic imaging modality soon after. Since then many imaging modalities were invented, with computed tomography, ultrasound, magnetic resonance imaging, and positron emission tomography among the commonly used, and more and more complex imaging procedures have been developed. Image information plays a crucial role in decision making at many stages in the patient care process, including detection, characterization, staging, treatment response assessment, monitoring of disease recurrence, as well as guiding interventional procedures, surgeries, and radiation therapy.

Cherney.K(2022)[8] Magnetic resonance imaging (MRI) has transformed the field of neurosurgery by improving the accuracy of diagnostic imaging, integration into presurgical planning, and moving into the operating room (OR) itself. For surgical applications, most innovations have occurred within the realm of stereotactic neurosurgery. There has been some controversy regarding credit for the discovery that the physical properties of nuclear magnetic resonance could be used to image living tissue. In 2003 the Nobel Prize in Medicine was awarded to Drs. Paul Lauterbur and Peter Mansfield for this work. Dr. Raymond Damadian publicly objected to his exclusion despite his pioneering efforts in the field [1,2]. Be that as it may, by the mid 1980s MRI had become a commercially available technology whose benefits were obvious. These included much better soft tissue imaging of the central nervous system, multiplanar views, and the avoidance of ionizing radiation. Within short order MRI was adapted to image guided surgery and is now the mainstay of this approach in the developed world.

Hao Ren et.al (2021)[6] With the rapid development of AI techniques, Computer-aided Diagnosis has attracted much attention and has been successfully deployed in many applications of health care and medical diagnosis. For some specific tasks, the learning-based system can compare with or even outperform human experts' performance. The impressive performance owes to the excellent expressiveness and scalability of the neural networks, although the models' intuition usually cannot be represented explicitly. Interpretability is, however, very important, even the same as the diagnosis precision, for computer-aided diagnosis. To fill this gap, our approach is intuitive to detect pneumonia interpretably. We first build a large dataset of community-acquired pneumonia consisting of 35389 cases (distinguished from nosocomial pneumonia) based on actual medical records. Second, we train a prediction model with the chest X-ray images in our dataset, capable of precisely detecting pneumonia. Third, we propose an intuitive approach to combine neural networks with an explainable model such as the Bayesian Network. The experiment result shows that our proposal further improves the performance by using multi-source data and provides intuitive explanations for the diagnosis results.

Jianpeng Zhang,et.al (2021)[7] Clusters of viral pneumonia occurrences over a short period may be a harbinger of an outbreak or pandemic. Rapid and accurate detection of viral pneumonia using chest X-rays can be of significant value for large-scale screening and epidemic prevention, particularly when other more sophisticated imaging modalities are not readily accessible. However, the emergence of novel mutated viruses causes a substantial dataset shift, which can greatly limit the performance of classification-based approaches. In this paper,we formulate the task of differentiating viral pneumonia from non-viral pneumonia and healthy controls into a one-class classification-based anomaly detection problem. We therefore propose the confidence-aware anomaly detection (CAAD) model, which consists of a shared feature extractor, an anomaly detection module, and a confidence prediction module. If the anomaly score produced by the anomaly detection module is large enough, or the confidence score estimated by the confidence prediction module is small enough, the input will be accepted as an anomaly case (i.e., viral pneumonia). The major advantage of our approach over binary classification is that we avoid modeling individual viral pneumonia classes explicitly and treat.

Miotto.R et.al (2018)[12] Health care is coming to a new era where the abundant biomedical data are playing more and more important roles. In this context, for example, precision medicine attempts to 'ensure that the right treatment is delivered to the right patient at the right time' by taking into account several aspects of patient's data, including variability in molecular traits, environment, electronic health records (EHRs) and lifestyle [1–3]. The large availability of biomedical data brings tremendous opportunities and challenges to health care research. In particular, exploring the associations among all the different pieces of information in these data sets is a fundamental problem to develop reliable medical tools based on data-driven approaches and machine learning. To this aim, previous works tried to link multiple data sources to build joint knowledge bases that could be used for predictive analysis and discovery [4–6]. Although existing models demonstrate great promises (e.g. [7–11]), predictive tools based on machine learning techniques have not been widely applied in medicine [12].

Mohammad Yaseliani et.al(2022)[10]Pneumonia is an acute respiratory infection that has led to significant deaths of people worldwide. This lung disease is more common in people older than 65 and children under five years old. Although the treatment of pneumonia can be challenging, it can be prevented by early diagnosis using Computer-Aided Diagnosis (CAD) systems. Chest X-Rays (CXRs) are currently the primary imaging tool for detection of pneumonia, which are widely used by radiologists. While the standard approach of detecting pneumonia is based on clinicians' decisions, various Deep Learning (DL) methods have been developed for detection of pneumonia considering CAD system. In this regard, a novel hybrid Convolutional Neural Network (CNN) model is proposed using three classification approaches. In the first classification approach, Fully-Connected

(FC) layers are utilized for the classification of CXR images. This model is trained for several epochs and the weights that result in the highest classification accuracy are saved. In the second classification approach, the trained optimized weights are utilized to extract the most representative CXR image features and Machine Learning (ML) classifiers are employed to classify the images. In the third classification approach, an ensemble of the proposed classifiers is created to classify CXR images. The results suggest that the proposed ensemble classifier using Support Vector Machine (SVM) with Radial Basis Function (RBF) and Logistic Regression (LR) classifiers has the best performance with 98.55% accuracy. Ultimately, this model is deployed to create a web-based CAD system to assist radiologists in pneumonia detection with a significant accuracy.

Nahiduzzaman.Md et.al (2020) [9] In this era of COVID19, proper diagnosis and treatment of pneumonia are very important. Chest X-Ray (CXR) image analysis plays a vital role in the reliable diagnosis of pneumonia. An experienced radiologist is required for this. However, even for an experienced radiographer, it is quite challenging and time-consuming to diagnose accurately due to the fuzziness of CXR images. Also, identification can be erroneous due to the involvement of human judgement. Hence, an authentic and automated system can play an important role here. In this era of cutting-edge technology, deep learning (DL) is highly used in every sector. There are several existing methods to diagnose pneumonia but they have accuracy problems. In this study, an automatic pneumonia detection system has been proposed by applying the extreme learning machine (ELM) on the Kaggle CXR images (Pneumonia). Three models have been studied: classification using extreme learning machine (ELM), ELM with a hybrid convolutional neural network-principal component analysis (CNN-PCA) based feature extraction, and CNN-PCA-ELM with the CXR images which are contrast-enhanced by contrast limited adaptive histogram equalization (CLAHE).

Suguna.S et.al (2017)[11] A study in 2016 found that human beings are collectively generated data more than ten exabytes, or 5x1018 bytes from various sources (Lyman and Varian 2003). Exploratory Data Analysis (EDA) is a method to analyze data using advanced techniques to expose hidden structure, enhances the insight into a given dataset, identifies the anomalies and builds parsimonious models to test the underlying assumptions. Exploratory Data Analysis (EDA) is classified into Graphical or non-graphical and Univariate or multivariate Univariate data consider one data column at a time while multivariate method considers more than two variables while analyzing. The diagnostic methods of diseases are of two types namely, Invasive and Non-invasive.

Sun.Y et.al (2019) [15] Pneumonia results in significant morbidity and mortality worldwide. However, chest radiography may not be accessible in primary care setting. We aimed to evaluate clinical features and its diagnostic value to identify pneumonia among adults in primary care settings. There is a need to improve the diagnosis accuracy. In this work, an efficient model for the detection of pneumonia trained on digital chest X-ray images is proposed, which could aid the radiologists in their decision making process Three academic databases were searched and included studies that assessed clinical predictors of pneumonia, adults without serious illness, have CXR and have conducted in primary care settings. Chest X-rays are primarily used for the diagnosis of this disease. However, even for a trained radiologist, it is a challenging task to examine chest X-rays. There is a need to improve the diagnosis accuracy. In this work, an efficient model for the detection of pneumonia trained on digital chest X-ray images is proposed, which could aid the radiologists in their decision making process. A novel approach based on a weighted.We calculated sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio of each index test and the pool estimates for index tests. We identified 2,397 articles, of which 13 articles were included. In our meta-analysis, clinical features with the best pooled positive likelihood ratios were respiratory rate ≥20min−1 (3.47; 1.46–7.23), temperature ≥38°C (3.21; 2.36–4.23), pulse rate >100min−1 (2.79; 1.71–4.33), and crackles (2.42; 1.19–4.69). Laboratory testing showed highest pooled positive likelihood ratios with PCT >0.25ng/ml (7.61; 3.28–15.1) and CRP>20mg/l (3.76; 2.3–5.91). Cough, pyrexia, tachycardia, tachypnea, and crackles are limited as a single predictor for diagnosis of radiographic pneumonia among adults. Development of clinical decision rule that combine these clinical features together with molecular biomarkers may further increase overall accuracy for diagnosis of radiographic pneumonia among adults in primary care setting.

Torres.A et.al (2021)[1] Pneumonia is a common acute respiratory infection that affects the alveoli and distal bronchial tree of the lungs. The disease is broadly divided into community-acquired pneumonia (CAP) or hospital-acquired pneumonia (HAP, which includes ventilation-associated pneumonia (VAP)) (Box 1). Aspiration pneumonia represents 5–15% of all cases of CAP; however, its prevalence amongst patients with HAP is not known . The lack of robust diagnostic criteria for aspiration pneumonia may explain why the true burden of this type of pneumonia remains unknown1 . The causative microorganisms for CAP and HAP differ substantially. The most common causal microorganisms in CAP are Streptococcus pneumoniae, respiratory viruses, Haemophilus influenzae and other bacteria such as Mycoplasma pneumoniae and Legionella pneumophila. Conversely, the most frequent microorganisms in HAP are Staphylococcus aureus (including both methicillin-susceptible S. aureus (MSSA) and methicillin-resistant S. aureus (MRSA)), Enterobacterales, non-fermenting gram-negative bacilli (for example, Pseudomonas aeruginosa), and Acinetobacter spp.2,3 . In health-care-associated pneumonia (HCAP), owing to patient risk factors, the microbial aetiology is more similar to that in HAP than to that in CAP. However, difficulties in standardizing risk factors for this population, coupled with the heterogeneity of post-hospital health care worldwide, suggest that the concept of HCAP has little usefulness, and indeed, HCAP was not included in recent guidelines for CAP and HAP3–5 .

Vinitha S et.al (2018)[13] With the advance of big data analytics equipment, more devotion has been paid to disease expectation from the perception of big data inquiry, various explores have been conducted by choosing the features mechanically from a large number of data to improve the truth of menace classification rather than the formerly selected physiognomies. those prevailing work mostly measured structured data.

Wasif Khan et.al (2021) [4]Chest radiography is a significant diagnostic tool used to detect diseases afflicting the chest. The automatic detection techniques associated with computer vision are being adopted in medical imaging research. Over the last decade, several remarkable advancements have been made in the field of medical diagnostics with the application of deep learning techniques. Various automated systems have been proposed for the rapid detection of pneumonia from chest X-rays. Although several algorithms are currently available for pneumonia detection, a detailed review summarizing the literature and offering guidelines for medical practitioners is lacking. This study will help practitioners to select the most effective and efficient methods from a real-time perspective, review the available datasets, and understand the results obtained in this domain. It will also present an overview of the literature on intelligent pneumonia identification from chest X-rays. The usability, goodness factors, and computational complexities of the algorithms employed detection.

CHAPTER 3

SYSTEM DESIGN

In this chapter, the various UML diagrams for Pneumonia Detection Using Chest X-Ray Images in Python are represented and the various functionalities are explained.

3.1 UNIFIED MODELING LANGUAGE

Unified Modeling Language (UML) is a standardized modeling language enabling developers to specify, visualize, construct and document artifacts of a software system. Thus, UML makes these artifacts scalable, secure, and robust in execution. It uses graphic notation to create visual models of software systems. UML is designed to enable users to develop an expressive, ready-to-use visual modeling language. In addition, it supports high-level development concepts such as frameworks, patterns, and collaborations. Some of the UML diagrams are discussed.

3.1.1 Use Case Diagram of Pneumonia Prediction Using Algorithms

Use case diagrams are considered for high-level requirement analysis of a system. So, when the requirements of a system are analyzed, the functionalities are captured in use cases. So, it can be said that use cases are nothing but system functionalities written in an organized manner. Now the second things which are relevant to the use cases are the actors. Actors can be defined as something that interacts with the system. The actors can be human users, some internal applications, or maybe some external applications. Use case diagrams are used to gather the requirements of a system including internal and external influences. These requirements are mostly design requirements .System, actors, use cases and relationships. Then, arrange them visually in a way that makes sense and will allow you to see immediately the connections between them. An actor can be a person, an organization, or another system. Use case diagrams show the expected behavior of the system.

They don't show the order in which steps are performed.The Use Case diagram of the project consists of all the various aspects a normal use case diagram requires. This use case diagram shows how, from starting the model flows from one step to another, like entering the dataset into the system .

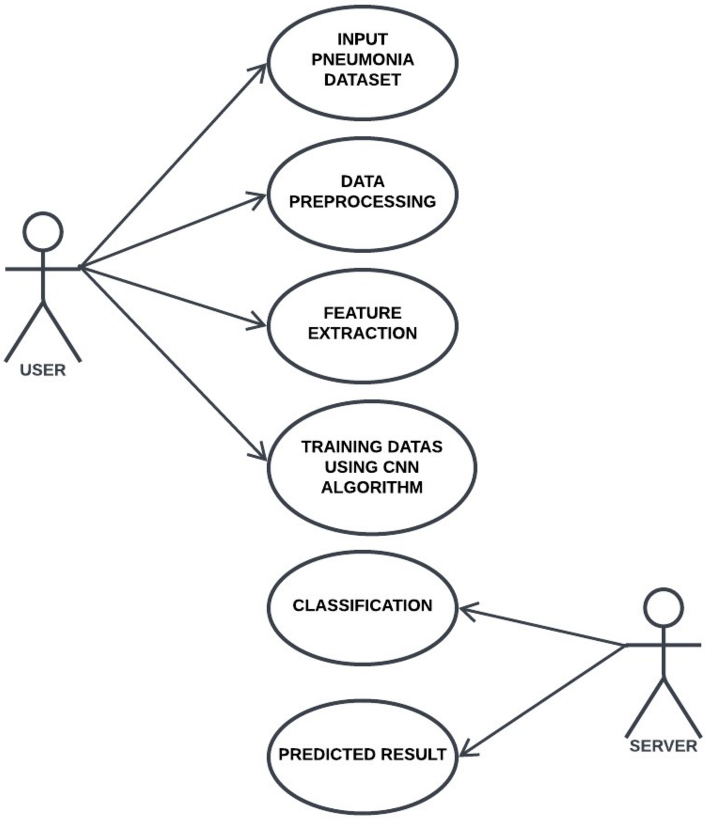

Figure 3.1 Use case diagram of Pneumonia Prediction Using Algorithms

Figure 3.1 shows that the functionalities are to be represented as a use case in the representation.Then entering all the information and all other general information that goes into the system, compares with the classification model and if true it provides the appropriate results otherwise it shows the details where the information is wrong. Here the use case diagram of all the entities is linked to each other when the user gets started with the system.As how the user inputs the information as the image input and it shows the output as the predictions from which we can determine best model.

3.1.2 Sequence Diagram of Pneumonia Prediction Using Algorithms

Figure 3.2 shows that UML sequence diagrams model the flow of logic within the system in a visual manner, enabling both documents and validating the logic, and are commonly used for both analysis and design purposes.

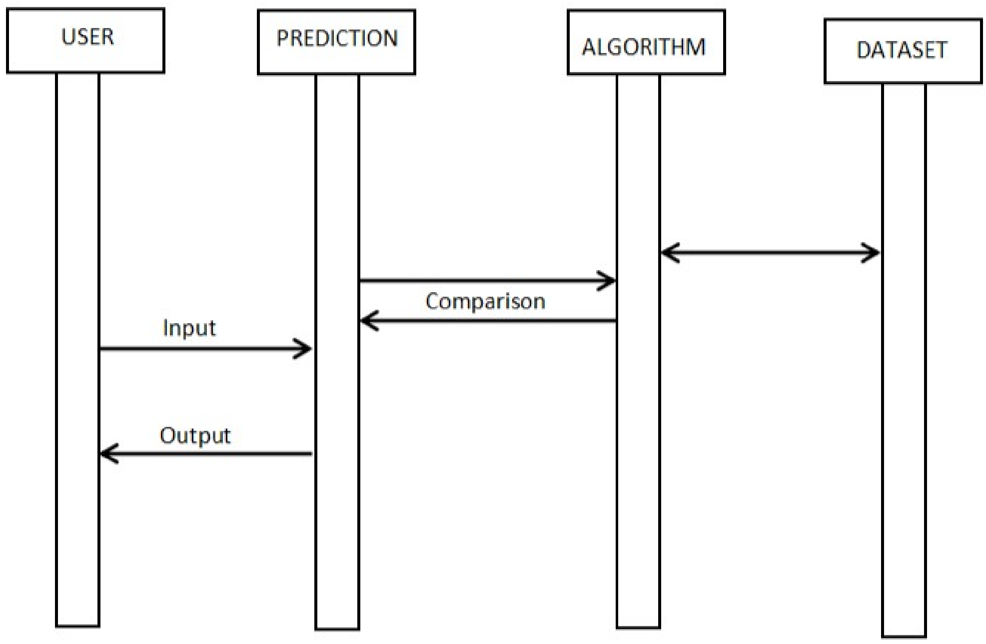

Figure 3.2 Sequence diagram of Pneumonia Prediction Using Algorithms

The various actions that take place in the application in the correct sequence are shown in Figure 3.2 Sequence diagrams are the most popular UML for dynamic modeling. This sequence diagram shows how starting the model flows from one step to another, like entering the dataset into the system and then entering all the information and all other general information that goes into the system, comparing with the classification model and if true provides the appropriate results . Here the sequence of all the entities is linked to each other and the user gets started with the system. Other dynamic modeling techniques include activity diagramming, communication diagramming, timing diagramming, and interaction overview diagramming.

3.1.3 Activity Diagram of Pneumonia Prediction Using Algorithms

Figure 3.3 shows that activity is a particular operation of the system. An activity diagram is suitable for modeling the activity flow of the system.

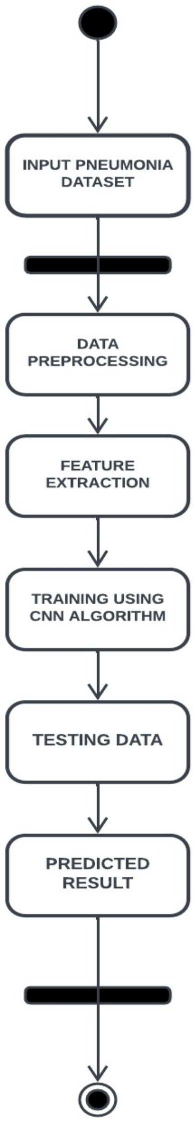

Figure 3.3 Activity Diagram of Pneumonia Prediction Using Algorithms

Activity diagrams are not only used for visualizing the dynamic nature of a system but they are also used to construct the executable system by using forward and reverse engineering techniques. An activity diagram is suitable for modeling the activity flow of the system. It does not show any message flow from one activity to another. An activity diagram is sometimes considered a flow chart Although the diagrams look like a flow chart but is not. It shows different flows like parallel, branched, concurrent and single. Fig 3.3 shows the activity diagram of the developed application. In our project, activity diagram flow starts from collecting datasets, cleaning and minimization, and training using a supervised machine learning algorithm and testing using the user input.

3.1.4 Collaboration Diagram of Pneumonia Prediction Using Algorithms

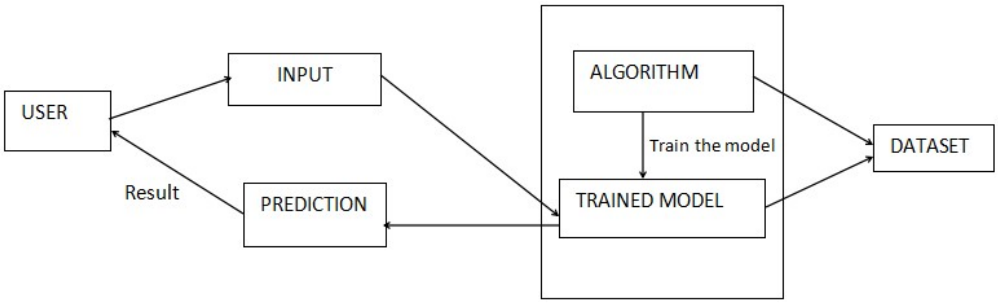

Figure 3.4 Collaboration Diagram of Pneumonia Prediction Using Algorithms

Figure 3.4 is a collaboration diagram, also known as a communication diagram, is an illustration of the relationships and interactions among software objects in the Unified Modelling Language (UML). These diagrams can be used to portray the dynamic behavior of a particular use case and define the role of each object.

Here this diagram shows how all the models are connected to show the correct result starting from the user using those datasets into the system and then applying all the information in order to get the accurate result, then the system evaluates the user-entered information. Finally gives the correct result.

CHAPTER 4
SYSTEM ARCHITECTURE

In this chapter, the System Architecture for Pneumonia Detection using Algorithms using Machine Learning is represented and the modules are explained.

4.1 SYSTEM ARCHITECTURE

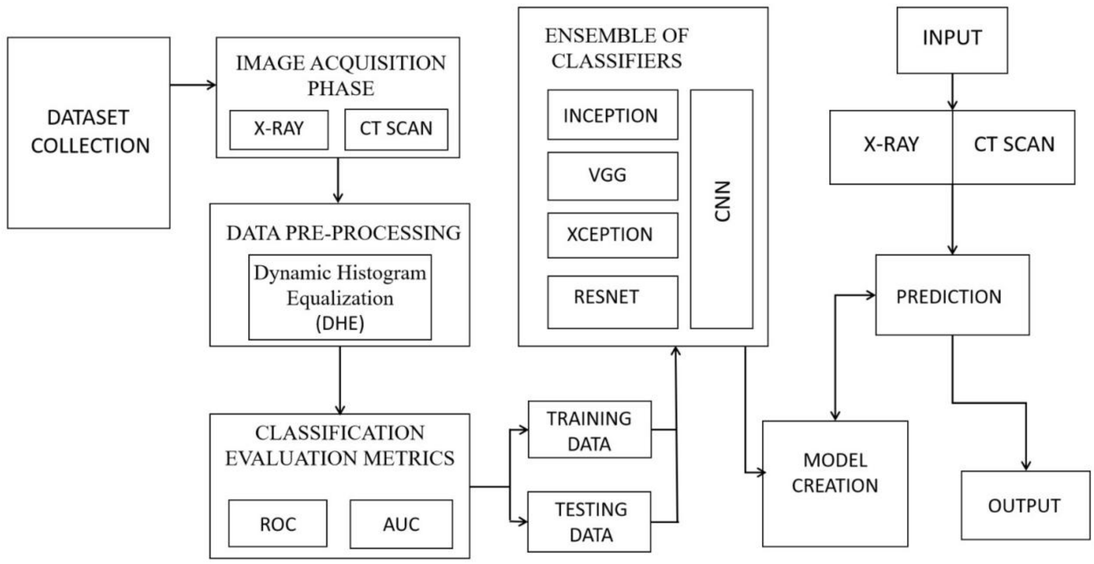

Figure 4.1 System Architecture

4.1.1 ARCHITECTURE DESCRIPTION

In figure 4.1 the architecture diagram is a graphical representation of a set of concepts, that are part of an architecture, including their principles, elements, and components. The diagram explains about the system software in the perception of an overview of the system. System architecture for Pneumonia detection using X-ray images involves the use of deep learning algorithms to classify X-ray images as normal or pneumonia-infected. The system will comprise of multiple components, including data pre-processing, model training, and testing, as well as deployment. This component involves preparing the dataset for training the deep learning model. The dataset will include X-ray images of patients with and without pneumonia. Data pre-processing techniques such as re-sizing, normalization, and augmentation will be used to improve the quality of the dataset. A CNN will be used as the deep learning model to detect pneumonia from X-ray images. The CNN will consist of multiple convolutional layers, pooling layers, and fully connected layers. The convolutional layers will extract features from the input X-ray images, while the fully connected layers will perform the classification task.The training of the CNN model will involve optimizing the parameters using back propagation and gradient descent. The model's performance will be evaluated using metrics such as accuracy, precision, recall, and F1 score. The testing of the model will involve using a separate set of X-ray images to evaluate its performance. The final component of the system architecture involves deploying the trained model in a production environment.The algorithm models used like inception,xception,vgg,resnet can be utilized. The model will be deployed as a web service, allowing healthcare professionals to upload X-ray images for diagnosis. The system will be integrated with a user interface that will display the diagnosis results to the user. Overall, the system architecture for pneumonia detection using X-ray images will involve the use of deep learning algorithms to accurately classify X-ray images as normal or pneumonia-infected. The system will be designed to be user-friendly and easily accessible to healthcare professionals. System architecture for Pneumonia detection using X-ray images involves the use of deep learning algorithms to classify X-ray images as normal or pneumonia-infected. The system will comprise of multiple components, including data pre-processing, model training, and testing, as well as deployment.Then when the input is given the system checks it with the trained model and gives the prediction with the probabilities.From which it can be determined the output.

CHAPTER 5

SYSTEM IMPLEMENTATION

In this chapter, the System Implementation for the Detection And Comparison Of Algorithm For Prediction Of Pneumonia is explained in detail.

5.1 MODULE DESCRIPTION

5.1.1 Image Acquisition Phase

The first step is to acquire images. To produce a classification model, the computer needs to learn by example. The computer needs to view many images to recognize an object. Other types of data, such as time series data and voice data, can also be used to train deep learning models. In the context of the work surveyed in this paper, the relevant data required to detect lung disease will be images. Images that could be used include chest X-ray, CT scan image. The output of this step is images that will later be used to train the model.

5.1.2 Data Pre-Processing

Image pre-processing is a very common and beneficial technique in the deep learning process and it not only could enlarge the quantity of the original dataset but also enrich the information implicit in the dataset. As previously mentioned, we utilized an effective image enhancement method named Dynamic Histogram Equalization (DHE) to improve the quality of images before they were inputted into the CNN model. Histogram Equalization (HE), which denotes mapping from the initial narrow pixel levels to a wider extent and improves image enhancement, has been widely used in image processing. The HE technique means to convert the gray levels of an image by using cumulative effort function globally, yet always brings about the problem that elaboration information in images is damaged, leading to awful image quality. This popular image contrast enhancement method could enhance image contrast effectively in many aspects, like MRI, X-rays, and CT.

5.1.3 Classification Evaluation Metrics

In this subsection, several evaluation metrics, accuracy, precision, recall, F1 score and so on, are described. According to the outputs of model, four indices, True Positive, True Negative, False Positive, False Negative, are used to analyze and identify the performance of model. The True Positive means that the chest X-ray images, which suffer from pneumonia, are signed as pneumonia as well by the model. The True Negative means if the chest X-ray images do not show pneumonia as well as the model predicts. The precision rate was always used to estimate how much the number of images that are truly pneumonia accounted for in the total number examples, which are classified as positive for pneumonia. That is, the pneumonia images must be identified in practical clinical diagnoses and hence, the precision rate is especially important. In most cases, the higher the precision rate gets, the lower the recall rate is. Thus, F1 score rate is widely considered as a proper criterion. In addition, the Receiver Operating Characteristic (ROC) and AUC are calculated to compare the performance of different models.

5.1.4 Ensemble Of Classifiers

When more than one classifier is combined to make a prediction, this is known as ensemble classification. Ensemble decreases the variance of predictions, therefore making predictions that are more accurate than any individual model. From work found in the literature, the ensemble techniques used include majority voting, probability score averaging and stacking. In majority voting, every model makes a prediction for each test instance, or, in other words, votes for a class label, and the final prediction is the label that received the most votes. An alternate version of majority voting is weighted majority voting, in which the votes of certain models are deemed more important than others. For example. In probability score averaging, the prediction scores of each model are added up and divided by the number of models involved. An alternate version of this is weighted averaging, where the prediction score of each model is multiplied by the weight, and then their average is calculated. Examples of works which used probability score averaging are found. In stacking ensemble, an algorithm receives the outputs of weaker models as input and tries to learn how to best combine the input predictions to provide a better output prediction.

5.2. ALGORITHM MODELS

5.2.1. Xception Model

Xception is a deep convolutional neural network architecture introduced by François Chollet in 2016. It stands for "Extreme Inception," and it is a variant of the Inception architecture.This is one of the models in system. The Xception model replaces the standard convolutional layers of the traditional CNN architecture with depth-wise separable convolutions. This type of convolutional layer performs two operations: first, it applies a depth-wise convolution, which computes a spatial convolution independently for each channel of the input tensor. Then, it applies a point-wise convolution, which performs a 1x1 convolution to combine the outputs of the depth-wise convolution.Which gives aboout an even increased performance for accuracy. The advantage of using depth-wise separable convolutions is that they significantly reduce the number of parameters in the model while maintaining a high level of accuracy.

5.2.2. Inception Model

The Inception model is a deep convolutional neural network architecture that was first introduced by Google researchers in 2014. It is designed to improve the accuracy of image classification tasks while keeping the computational cost relatively low.The name it gets from the given model The Inception model gets its name from the fact that it uses multiple "inception modules" in its architecture. These modules are composed of a series of convolutions of different sizes and are designed to capture features of different scales in an image. The main innovation of the Inception model is the use of so-called "1x1 convolutions." These are convolutions with a kernel size of 1x1 that are used to reduce the dimensionality of the feature maps before passing them to larger convolutions. This helps to reduce the computational cost of the model while improving its accuracy.As for another, key feature of the Inception model is its use of "parallel paths" in its architecture. By using multiple paths, each with a different set of convolutions, the model can capture different types of features and improve its accuracy.So,the result of this model can be a little more accurate with input since its introduction, the Inception model has been improved upon with the introduction of the Inception V2, Inception V3, and Inception V4 models, which have achieved state-of-the-art results on a variety of image classification benchmarks. The Inception model is widely used in applications such as image and video analysis, autonomous vehicles, and medical imaging.

5.2.3. VGG Model

The VGG model is a deep convolutional neural network architecture that was first introduced by a group of researchers from the University of Oxford in 2014. It is named after the Visual Geometry Group (VGG) at the university. The VGG model is characterized by its use of very small 3x3 convolutional filters. These filters are used repeatedly throughout the network, resulting in a deep architecture with many layers. The VGG model also uses max pooling layers to reduce the spatial dimensions of the feature maps. The VGG model is known for its simplicity and ease of implementation, and it has achieved state-of-the-art results on several image classification benchmarks. However, its main drawback is its high computational cost due to its large number of parameters. Since its introduction, the VGG model has been improved upon with the introduction of the VGG16 and VGG19 models, which have 16 and 19 layers, respectively. These models have achieved even better results on image classification tasks and are widely used in computer vision applications, including object detection, image segmentation, and style transfer.

5.2.4. RESNET Model

The ResNet (short for Residual Network) model is a deep convolutional neural network architecture that was first introduced by researchers from Microsoft Research in 2015. It is designed to address the vanishing gradient problem that can occur in very deep neural networks. The vanishing gradient problem occurs when gradients become very small as they are propagated backward through the network during training. This can make it difficult for the network to learn effectively, especially in very deep architectures. The ResNet model addresses this problem by introducing skip connections, also known as residual connections, which allow gradients to flow more easily through the network. These connections allow the network to learn residual functions that are easier to optimize, rather than trying to learn the desired mapping directly. The ResNet model is characterized by its use of "residual blocks," which contain multiple convolutional layers and skip connections. These blocks are repeated throughout the network, resulting in a very deep architecture with hundreds of layers. The vanishing gradient problem occurs when gradients become very small as they are propagated backward through the network during training. The ResNet model has achieved state-of-the-art results on several image classification benchmarks, and it has been widely used in computer vision applications, including object detection, image segmentation, and facial recognition. It is considered to be one of the most important breakthroughs in deep learning in recent years.

CHAPTER 6

RESULTS AND CODING

6.1 SAMPLE CODE

```
from flask import Flask, render_template, request, session, redirect,
url_for, flash

import os

from werkzeug.utils import secure_filename

from tensorflow.keras.models import load_model

import matplotlib.pyplot as plt

import cv2

import numpy as np

UPLOAD_FOLDER = './flask app/assets/images'

ALLOWED_EXTENSIONS = set(['png', 'jpg', 'jpeg', 'gif'])

app = Flask(__name__,static_url_path='/assets',

            static_folder='./flask app/assets',

            template_folder='./flask app')

app.config['UPLOAD_FOLDER'] = UPLOAD_FOLDER

@app.route('/')

def root():

  return render_template('index.html')

@app.route('/index.html')

def index():
```

```
    return render_template('index.html')
@app.route('/contact.html')
def contact():
    return render_template('contact.html')
@app.route('/news.html')
def news():
    return render_template('news.html')
@app.route('/about.html')
def about():
    return render_template('about.html')
@app.route('/faqs.html')
def faqs():
    return render_template('faqs.html')
@app.route('/prevention.html')
def prevention():
    return render_template('prevention.html')
@app.route('/upload.html')
def upload():
    return render_template('upload.html')
@app.route('/upload_chest.html')
def upload_chest():
    return render_template('upload_chest.html')
```

```
@app.route('/upload_ct.html')
def upload_ct():
  return render_template('upload_ct.html')
@app.route('/uploaded_chest', methods = ['POST', 'GET'])
def uploaded_chest():
  if request.method == 'POST':
      if 'file' not in request.files:
          flash('No file part')
          return redirect(request.url)
      file = request.files['file']
      if file.filename == '':
          flash('No selected file')
          return redirect(request.url)
      if file:
          file.save(os.path.join(app.config['UPLOAD_FOLDER'],
'upload_chest.jpg'))
  resnet_chest = load_model('models/resnet_chest.h5')
  vgg_chest = load_model('models/vgg_chest.h5')
  inception_chest = load_model('models/inceptionv3_chest.h5')
  xception_chest = load_model('models/xception_chest.h5')
  image = cv2.imread('./flask app/assets/images/upload_chest.jpg')
  image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
```

```
  image = cv2.resize(image,(224,224))
  image = np.array(image) / 255
  image = np.expand_dims(image, axis=0)
  resnet_pred = resnet_chest.predict(image)
  probability = resnet_pred[0]
  print("Resnet Predictions:")
  if probability[0] > 0.5:
    resnet_chest_pred = str('%.2f' % (probability[0]*100) + '% PNEUMONIA DETECTED')
  else:
    resnet_chest_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(resnet_chest_pred)
  with open('RESULT.txt', 'w') as f:
    f.write("Resnet Predictions:")
    f.write(resnet_chest_pred)
    f.write('\n')
  vgg_pred = vgg_chest.predict(image)
  probability = vgg_pred[0]
  print("VGG Predictions:")
  if probability[0] > 0.5:
    vgg_chest_pred = str('%.2f' % (probability[0]*100) + '% PNEUMONIA DETECTED')
```

```
  else:
    vgg_chest_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(vgg_chest_pred)
  with open('RESULT.txt', 'a') as f:
 f.write("VGG Predictions:")
    f.write(vgg_chest_pred)
    f.write('\n')
  inception_pred = inception_chest.predict(image)
  probability = inception_pred[0]
  print("Inception Predictions:")
  if probability[0] > 0.5:
    inception_chest_pred = str('%.2f' % (probability[0]*100) + '% PNEUMONIA DETECTED')
  else:
    inception_chest_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(inception_chest_pred)
  with open('RESULT.txt', 'a') as f:
    f.write("Inception Predictions:")
    f.write(inception_chest_pred)
    f.write('\n')
  xception_pred = xception_chest.predict(image)
```

```
  probability = xception_pred[0]
  print("Xception Predictions:")
  if probability[0] > 0.5:
    xception_chest_pred = str('%.2f' % (probability[0]*100) + '% PNEUMONIA DETECTED')
  else:
    xception_chest_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(xception_chest_pred)
  with open('RESULT.txt', 'a') as f:
     f.write("Xception Predictions:")
    f.write(xception_chest_pred)
    f.write('\n')
  return render_template('results_chest.html',resnet_chest_pred=resnet_chest_pred,vgg_chest_pred=vgg_chest_pred,inception_chest_pred=inxception_chest_pred,xception_chest_pred=xception_chest_pred)
@app.route('/uploaded_ct', methods = ['POST', 'GET'])
def uploaded_ct():
  if request.method == 'POST':
      if 'file' not in request.files:
          flash('No file part')
          return redirect(request.url)
```

```
        file = request.files['file']
       if file.filename == '':
           flash('No selected file')
           return redirect(request.url)
        if file:
           file.save(os.path.join(app.config['UPLOAD_FOLDER'],
'upload_ct.jpg'))
   resnet_ct = load_model('models/resnet_ct.h5')
   vgg_ct = load_model('models/vgg_ct.h5')
   inception_ct = load_model('models/inception_ct.h5')
   xception_ct = load_model('models/xception_ct.h5')
   image = cv2.imread('./flask app/assets/images/upload_ct.jpg')
   image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
   image = cv2.resize(image,(224,224))
   image = np.array(image) / 255
   image = np.expand_dims(image, axis=0)
   resnet_pred = resnet_ct.predict(image)
   probability = resnet_pred[0]
   print("Resnet Predictions:")
   if probability[0] > 0.5:
     resnet_ct_pred = str('%.2f' % (probability[0]*100) + '%
PNEUMONIA DETECTED')
```

```
  else:
    resnet_ct_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(resnet_ct_pred)
  with open('RESULT.txt', 'w') as f:
    f.write("Resnet Predictions:")
    f.write(resnet_ct_pred)
    f.write('\n')
  vgg_pred = vgg_ct.predict(image)
  probability = vgg_pred[0]
  print("VGG Predictions:")
  if probability[0] > 0.5:
    vgg_ct_pred = str('%.2f' % (probability[0]*100) + '% PNEUMONIA DETECTED')
  else:
    vgg_ct_pred = str('%.2f' % ((1-probability[0])*100) + '% PNEUMONIA NOT DETECTED')
  print(vgg_ct_pred)
  with open('RESULT.txt', 'a') as f:
    f.write("VGG Predictions:")
    f.write(vgg_ct_pred)
    f.write('\n')
```

```
  inception_pred = inception_ct.predict(image)
  probability = inception_pred[0]
  print("Inception Predictions:")
  if probability[0] > 0.5:
    inception_ct_pred = str('%.2f' % (probability[0]*100) + '%
PNEUMONIA DETECTED')
  else:
    inception_ct_pred = str('%.2f' % ((1-probability[0])*100) + '%
PNEUMONIA NOT DETECTED')
PNEUMONIA NOT DETECTED')
  print(xception_ct_pred)
  with open('RESULT.txt', 'a') as f:
    f.write("Xception Predictions:")
    f.write(xception_ct_pred)
    f.write('\n')
  return
render_template('results_ct.html',resnet_ct_pred=resnet_ct_pred,v
gg_ct_pred=vgg_ct_pred,inception_ct_pred=inception_ct_pred,xc
eption_ct_pred=xception_ct_pred)
if __name__ == '__main__':
  app.secret_key = ".."
  app.run()
```

6.2 SAMPLE SCREENSHOTS

6.2.1 JUPYTER NOTEBOOK

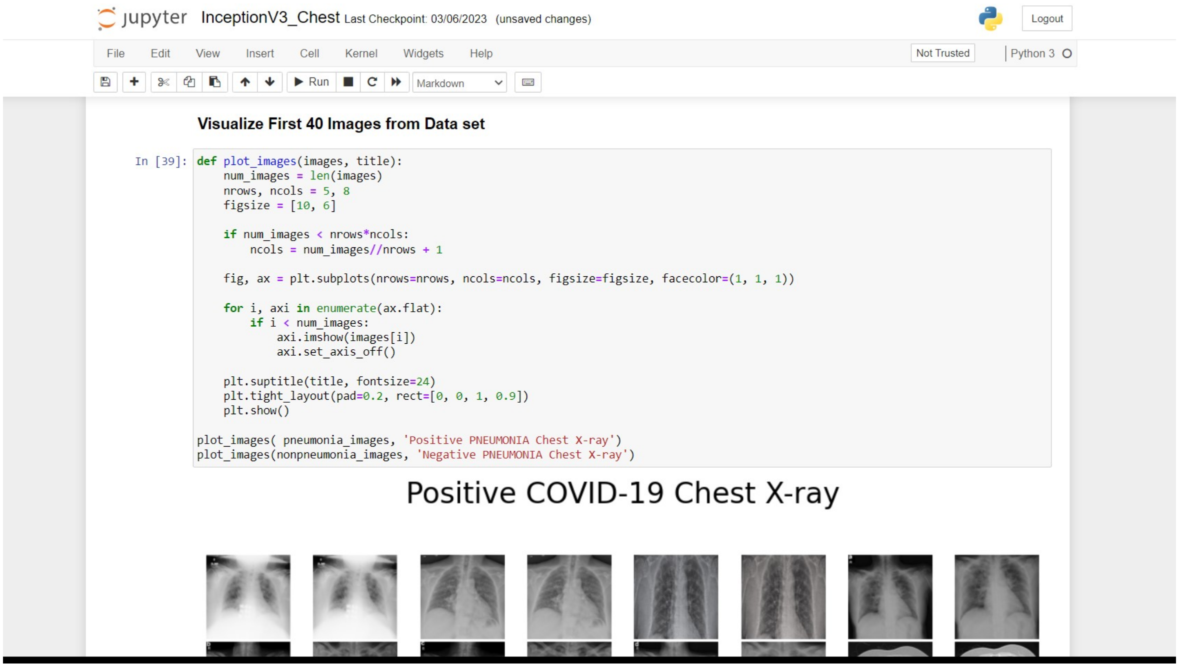

Figure 6.1 Screenshot of Jupyter Notebook

In figure 6.1, the screenshot for the jupyter notebook for the coding part of the system.It shows the positive covid-19 chest x-ray from the model which derives from the dataset collected.

6.2.2 PNEUMONIA DATASET

Figure 6.2 Screenshot of Pneumonia Dataset

This figure 6.2 shows the dataset included for this system which are the x-ray images of the pneumonia and non-pneumonia x-ray images.

6.2.3 RECEIVER OPERATING CHARACTERISTIC CURVE(ROC)

The receiver operating characteristic curve (ROC) for the different algorithm models used are given below as follows.

A. Xception Model

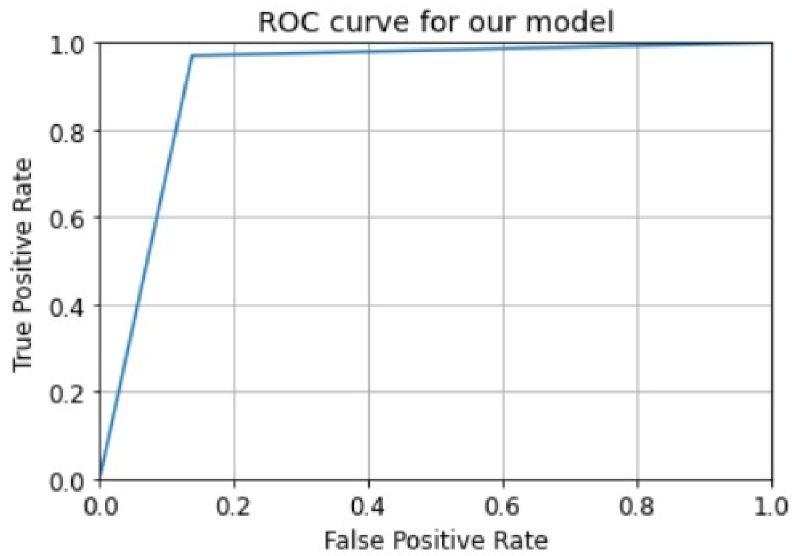

Figure 6.3 ROC curve of Xception Model using Chest X-Ray Images

In figure 6.3 shows the receiver operating curve of Xception algorithm model using Chest X-Ray Images from the dataset.

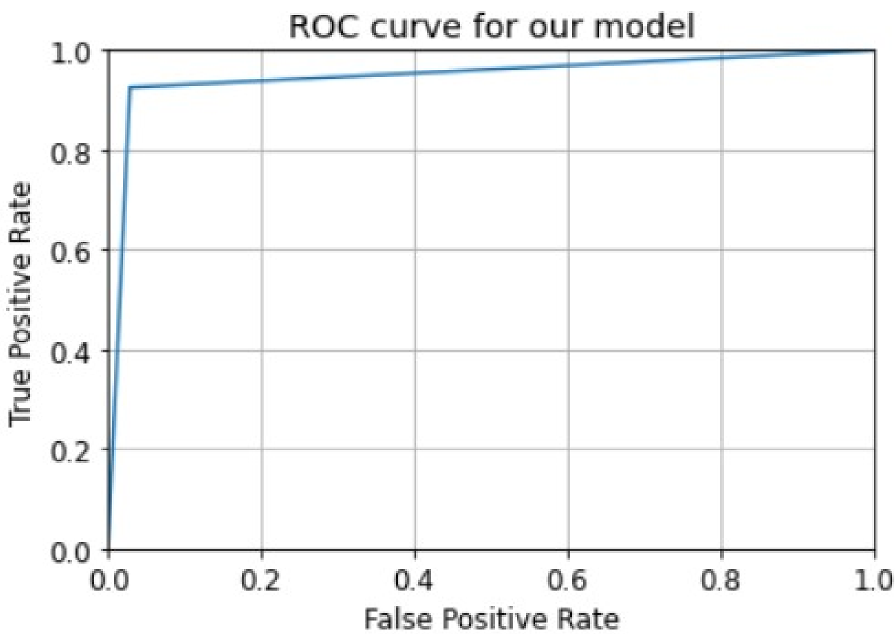

Figure 6.4 ROC curve of Xception Model using CT Scan

In figure 6.4 shows the receiver operating curve of Xception algorithm model using CT Scan from the dataset.

B. Inception Model

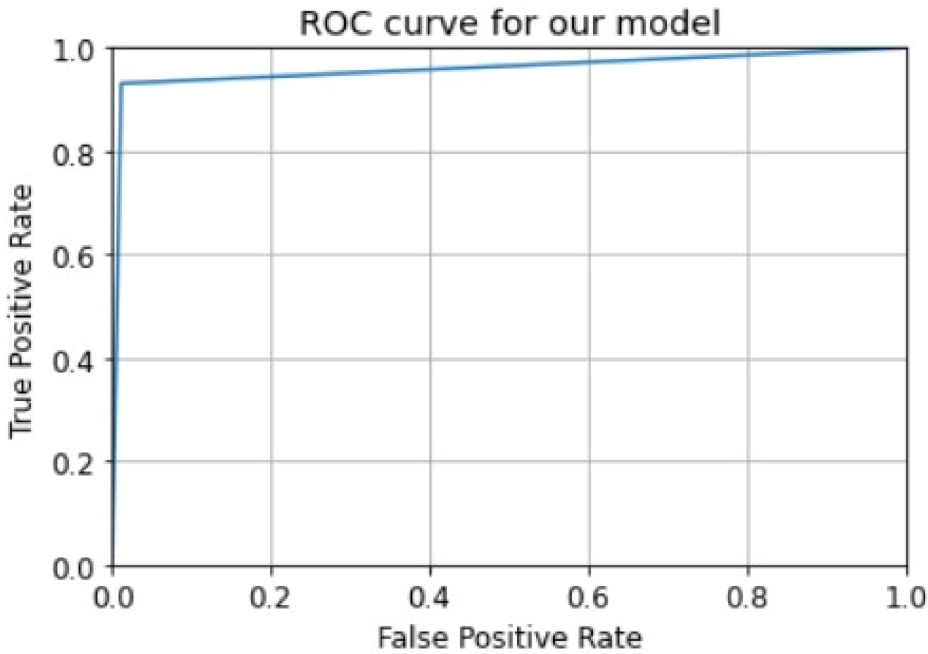

Figure 6.5 ROC curve of Inception Model using Chest X-Ray Images

In figure 6.5 shows the receiver operating curve of Inception algorithm model using Chest X-Ray Images from the dataset.

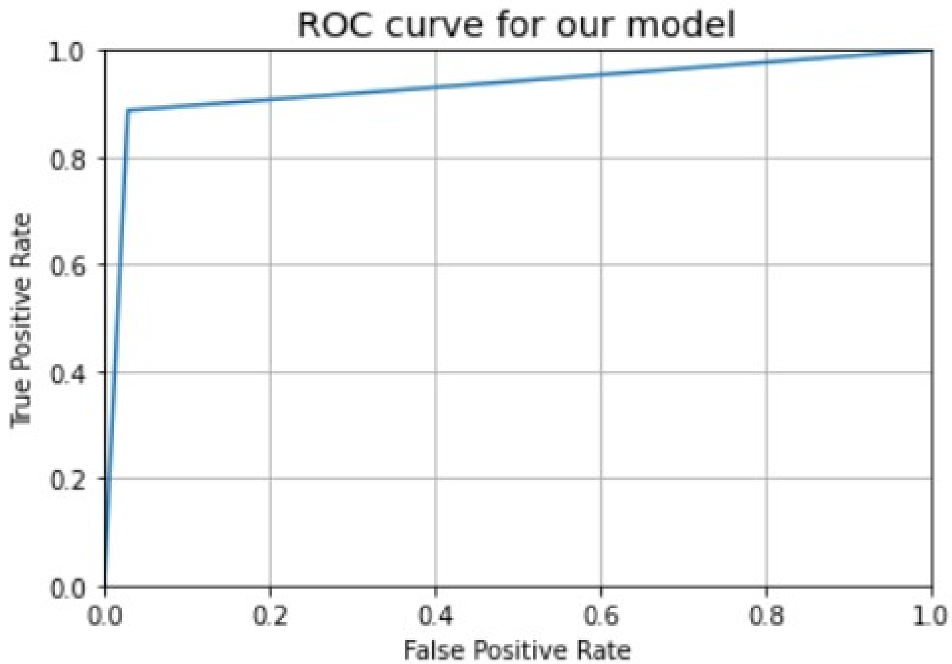

Figure 6.6 ROC curve of Inception Model using CT Scan

In figure 6.6 shows the receiver operating curve of Inception algorithm model using CT Scans from the dataset.

C.VGG Model

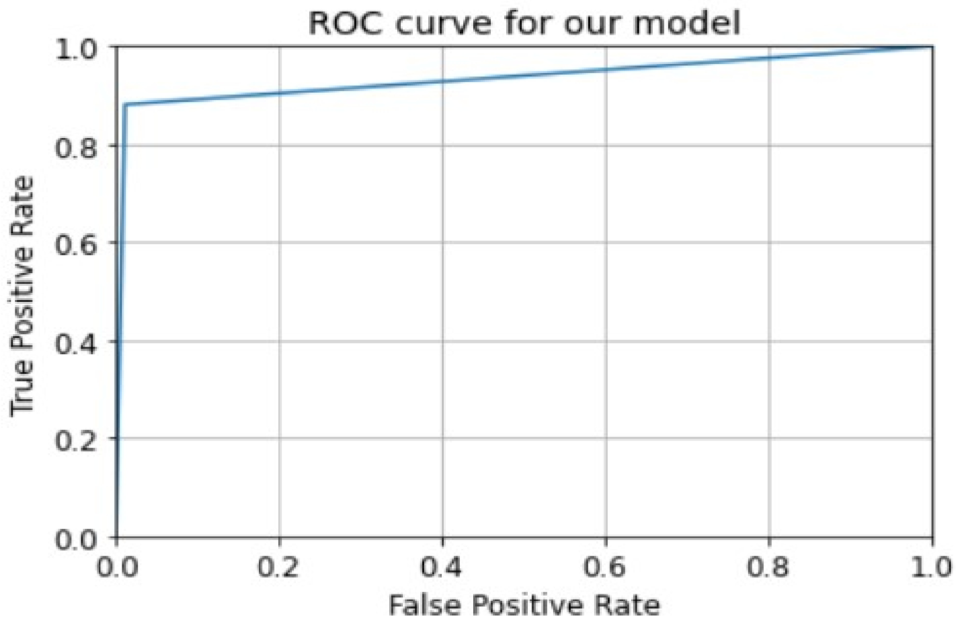

Figure 6.7 ROC curve of VGG Model using X-Ray Images

In figure 6.7 shows the receiver operating curve of VGG algorithm model using Chest X-Ray Images from the dataset.

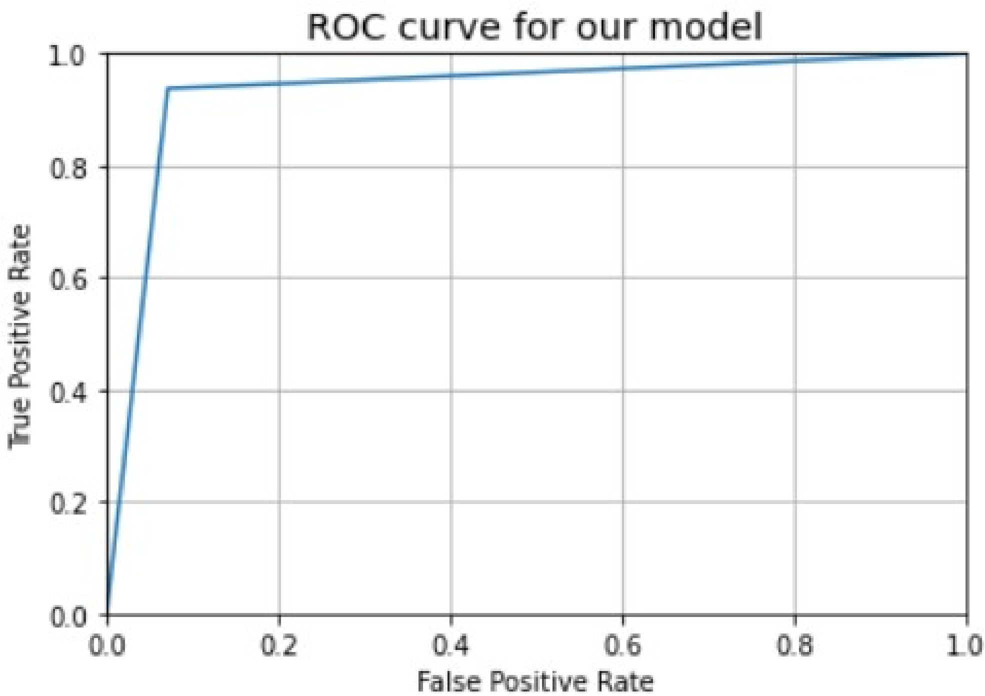

Figure 6.8 ROC curve of VGG Model using CT Scan

In figure 6.8 shows the receiver operating curve of VGG algorithm model using CT Scans from the dataset.

6.2.3.4 RESNET Model

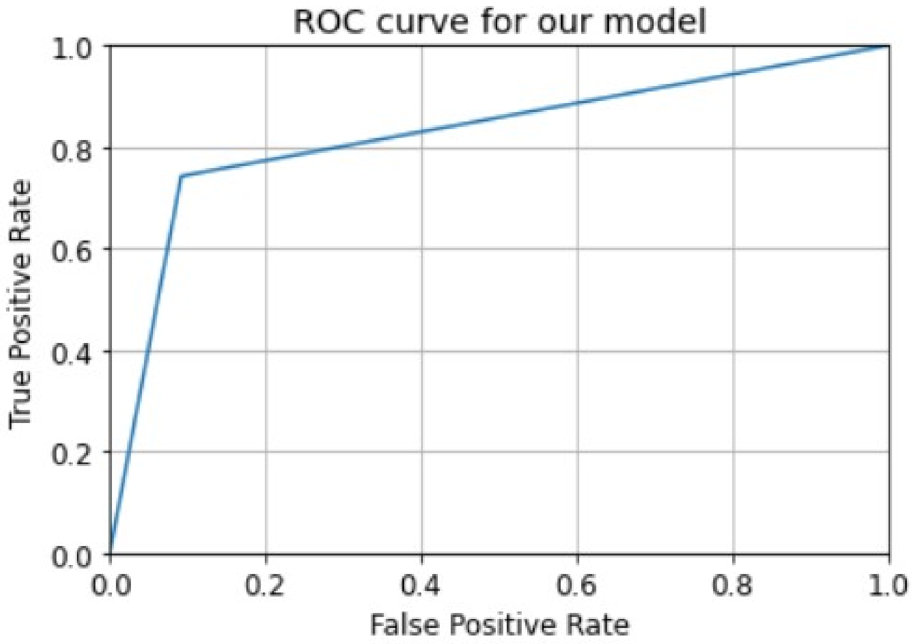

Figure 6.9 ROC curve of RESNET Model using Chest X-Ray Images

In figure 6.9 shows the receiver operating curve of RESNET algorithm model using Chest X-Ray Images from the dataset.

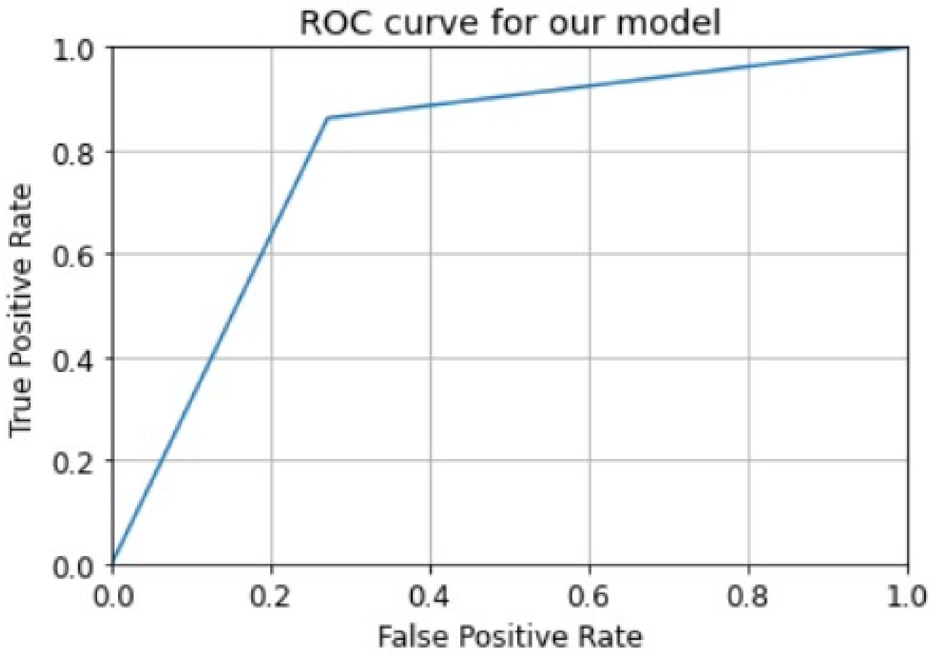

Figure 6.10 ROC curve of RESNET Model using CT Scan

In figure 6.10, it shows the receiver operating curve of RESNET algorithm model using Chest X-Ray Images from the dataset.

6.2.4 CLASSIFICATION REPORT

The classification report of different algorithm models used in this system is given below as follows.

A. Xception Model

	precision	recall	f1-score	support
0	0.96	0.86	0.91	87
1	0.89	0.97	0.93	101
accuracy			0.92	188
macro avg	0.93	0.92	0.92	188
weighted avg	0.92	0.92	0.92	188

Figure 6.11 Classification Report for Xception Model using Chest X-Ray Images

In figure 6.11, it shows the classification report of the Xception model using Chest X-Ray images from the dataset. Classification report for an Xception model using Chest X-Ray Images, we would need to have access to the ground truth labels for a dataset of Chest X-Ray Images, as well as the predictions made by the Xception model on that dataset.

	precision	recall	f1-score	support
0	0.92	0.97	0.94	70
1	0.97	0.93	0.95	80
accuracy			0.95	150
macro avg	0.95	0.95	0.95	150
weighted avg	0.95	0.95	0.95	150

Figure 6.12 Classification Report for Xception Model using CT Scan

In figure 6.12, it shows the classification report of the Xception model using CT Scans from the dataset. Classification report for an Xception model using CT scans, we would need to have access to the ground truth labels for a dataset of CT scans, as well as the predictions made by the Xception model on that dataset.

B. Inception Model

	precision	recall	f1-score	support
0	0.92	0.99	0.96	87
1	0.99	0.93	0.96	101
accuracy			0.96	188
macro avg	0.96	0.96	0.96	188
weighted avg	0.96	0.96	0.96	188

Figure 6.13 Classification Report for Inception Model using Chest X-Ray

In figure 6.13, it shows the classification report of the Iception model using Chest X-Ray images from the dataset. Classification report for an Inception model using Chest X-Ray Images, we would need to have access to the ground truth labels for a dataset of Chest X-Ray Images, as well as the predictions made by the Inception model on that dataset.

	precision	recall	f1-score	support
0	0.88	0.97	0.93	70
1	0.97	0.89	0.93	80
accuracy			0.93	150
macro avg	0.93	0.93	0.93	150
weighted avg	0.93	0.93	0.93	150

Figure 6.14 Classification Report for Inception Model using CT Scan

In figure 6.14, it shows the classification report of the Inception model using CT Scans from the dataset. Classification report for an Inception model using CT Scans, we would need to have access to the ground truth labels for a dataset of CT Scans, as well as the predictions made by the Inception model on that dataset.

C. VGG Model

	precision	recall	f1-score	support
0	0.88	0.99	0.93	87
1	0.99	0.88	0.93	101
accuracy			0.93	188
macro avg	0.93	0.93	0.93	188
weighted avg	0.94	0.93	0.93	188

Figure 6.15 Classification Report for VGG Model using Chest X-Ray

In figure 6.15, it shows the classification report of the VGG model using Chest X-Ray Images from the dataset. Classification report for an VGG model using Chest X-Ray Images, we would need to have access to the ground truth labels for a dataset of Chest X-Ray Images, as well as the predictions made by the VGG model on that dataset.

	precision	recall	f1-score	support
0	0.93	0.93	0.93	70
1	0.94	0.94	0.94	80
accuracy			0.93	150
macro avg	0.93	0.93	0.93	150
weighted avg	0.93	0.93	0.93	150

Figure 6.16 Classification Report for VGG Model using CT Scan

In figure 6.16, it shows the classification report of the VGG model using CT Scans from the dataset. Classification report for an VGG model using CT Scans, we would need to have access to the ground truth labels for a dataset of CT Scans, as well as the predictions made by the VGG model on that dataset.

D. RESNET Model

	precision	recall	f1-score	support
0	0.75	0.91	0.82	87
1	0.90	0.74	0.82	101
accuracy			0.82	188
macro avg	0.83	0.83	0.82	188
weighted avg	0.83	0.82	0.82	188

Figure 6.17 Classification Report for RESNET Model using Chest X-Ray

In figure 6.17, it shows the classification report of the RESNET model using Chest X-Ray Images from the dataset. Classification report for an RESNET model using Chest X-Ray Images, we would need to have access to the ground truth labels for a dataset of Chest X-Ray Images, as well as the predictions made by the RESNET model on that dataset.

	precision	recall	f1-score	support
0	0.82	0.73	0.77	70
1	0.78	0.86	0.82	80
accuracy			0.80	150
macro avg	0.80	0.80	0.80	150
weighted avg	0.80	0.80	0.80	150

Figure 6.18 Classification Report for RESNET Model using CT Scan

In figure 6.18, it shows the classification report of the RESNET model using CT Scans from the dataset. Classification report for an RESNET model using Chest X-Ray Images, we would need to have access to the ground truth labels for a dataset of CT Scans, as well as the predictions made by the RESNET model on that dataset.

6.2.5 COMPARISON OF ALGORITHMS

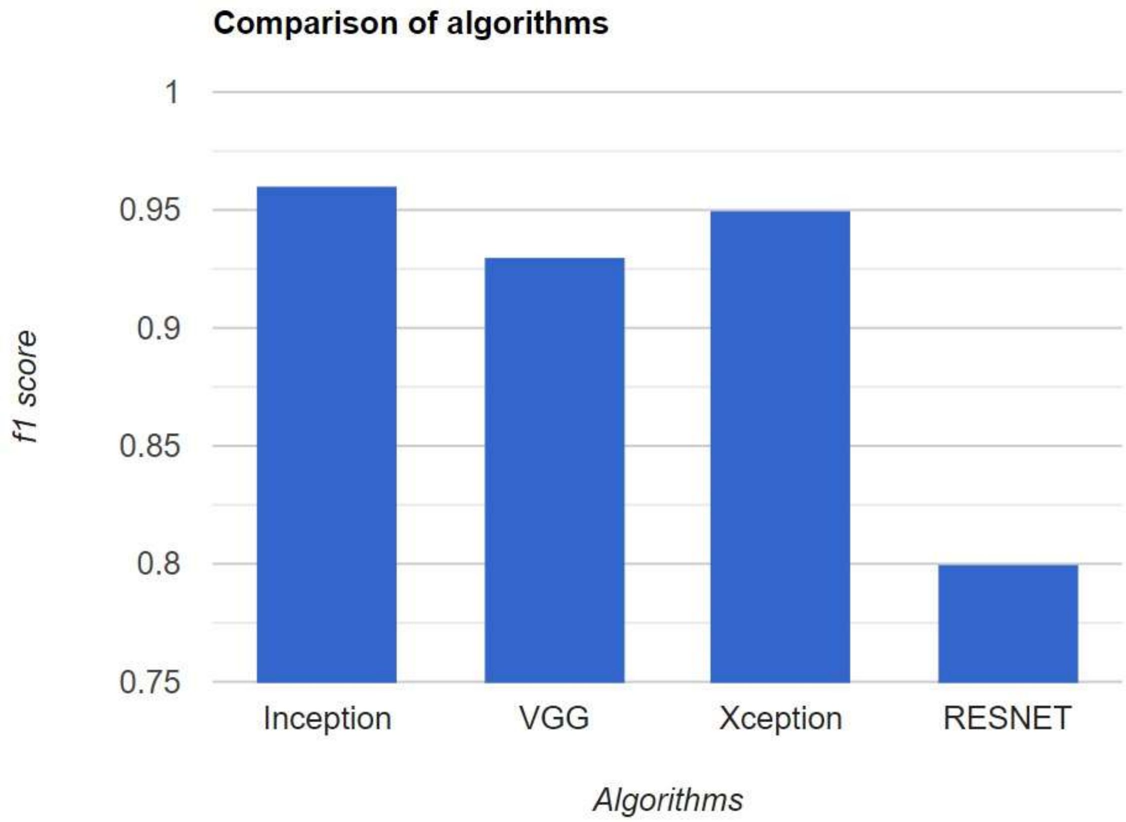

Figure 6.19 Comparison of Algorithm Models

In figure 6.19, the percentage of the algorithms have been compared and are represented in the bar graph as comparison of algorithms.

Various tests have been done were it was found that Inception is the most reliable model among the other three models with a f1 score of 0.96 and followed by Xception is 0.95, VGG which is 0.93 and RESNET which is the least favoured model which makes it totally apart from other three models with f1 score of 0.8.

Now it can be said that Inception Model is the most accurate algorithm model used to predict Pneumonia using Chest X-Ray Images and CT Scans from the datasets.

6.2.6 CONFUSION MATRIX

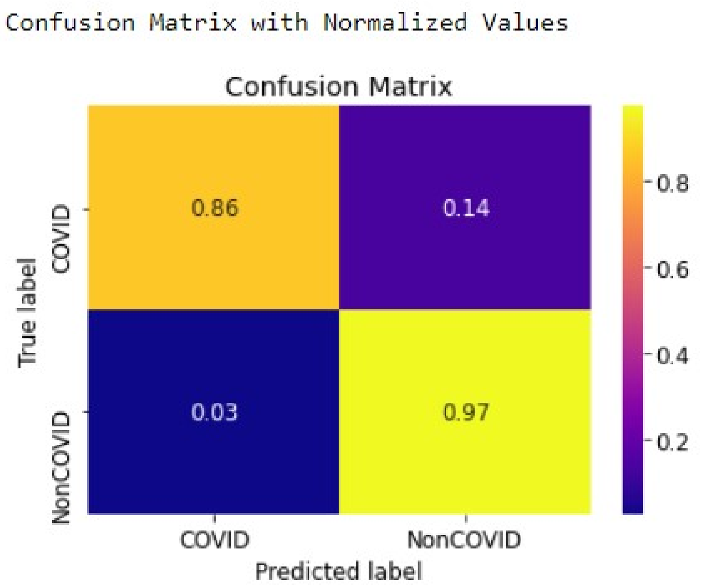

Figure 6.20 Confusion Matrix with Normalized Values using Chest X-ray images

In figure 6.20, it shows the confusion matrix with Normalized Values using Chest X-Ray Images from the dataset.

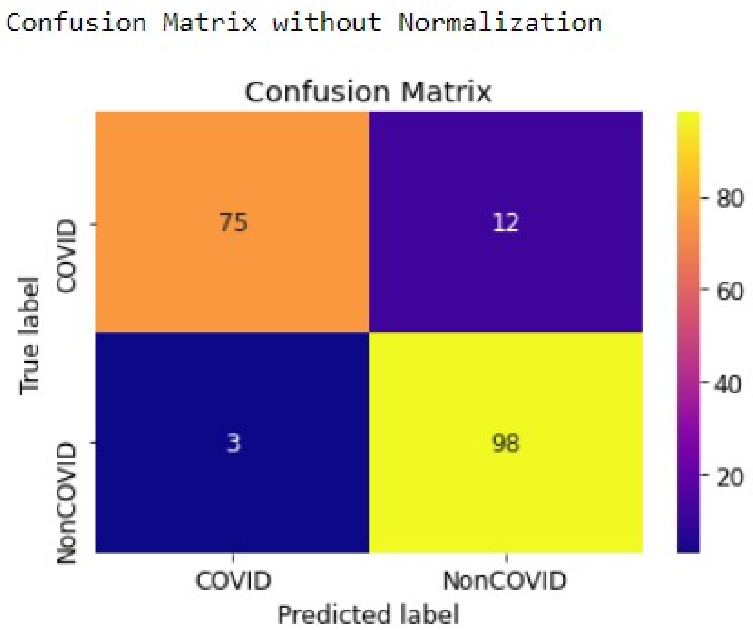

Figure 6.21 Confusion Matrix without Normalized Values using Chest X-Ray Images

In figure 6.21, it shows the confusion matrix without normalized values using Chest X-Ray Images from the dataset.

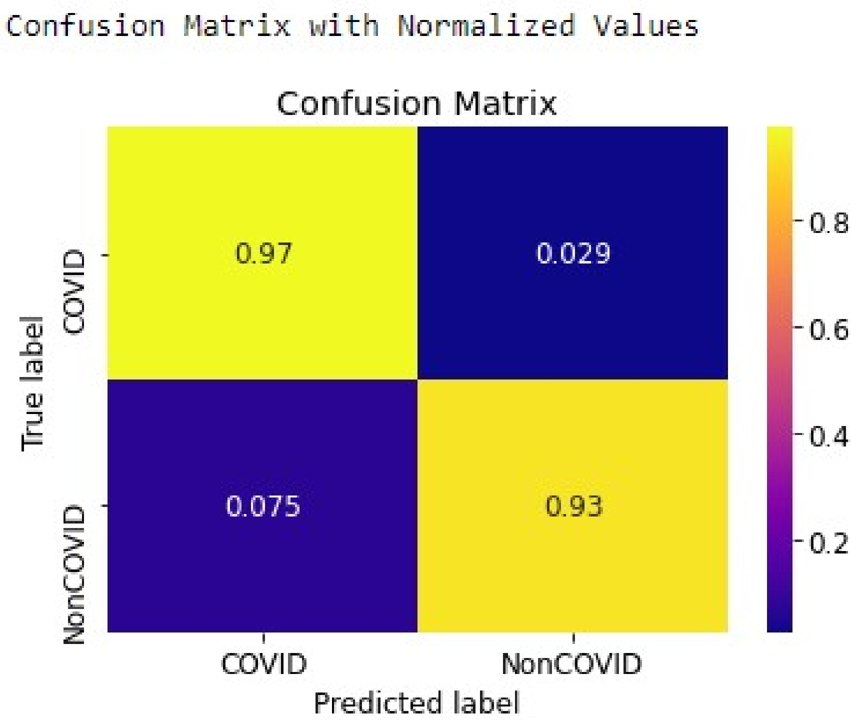

Figure 6.22 Confusion Matrix with Normalized Values with CT Scans

In figure 6.22, it shows the confusion matrix with normalized values using CT Scan from the dataset.

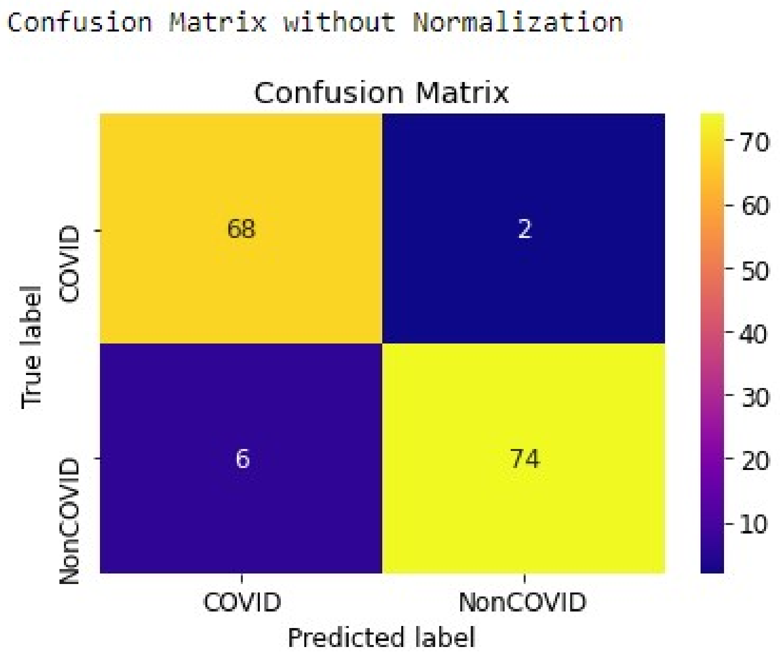

Figure 6.23 Confusion Matrix without Normalized Values with CT Scans

In figure 6.22, it shows the confusion matrix with normalized values using CT Scan from the dataset.

6.2.7 ACCURACY OF MODELS

A. Xception Model

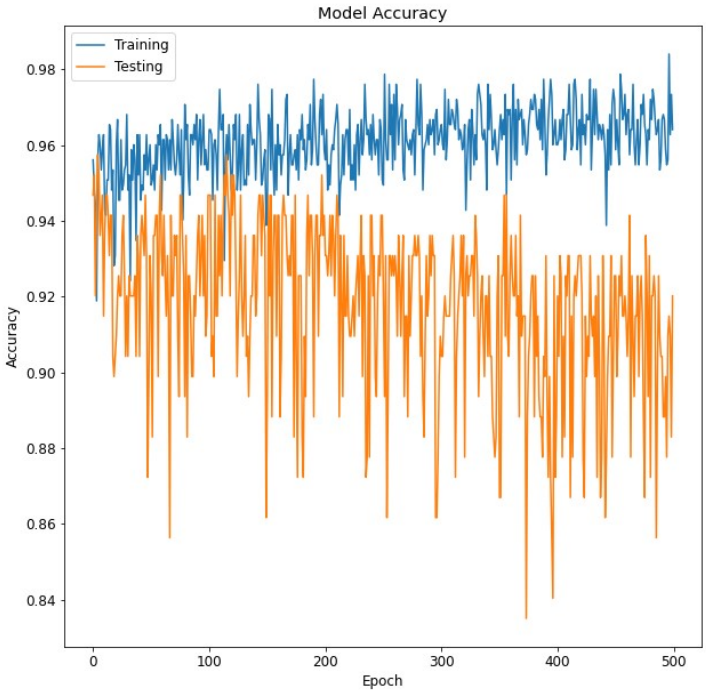

Figure 6.24 Accuracy Model of Xception Model using Chest X-Ray Images

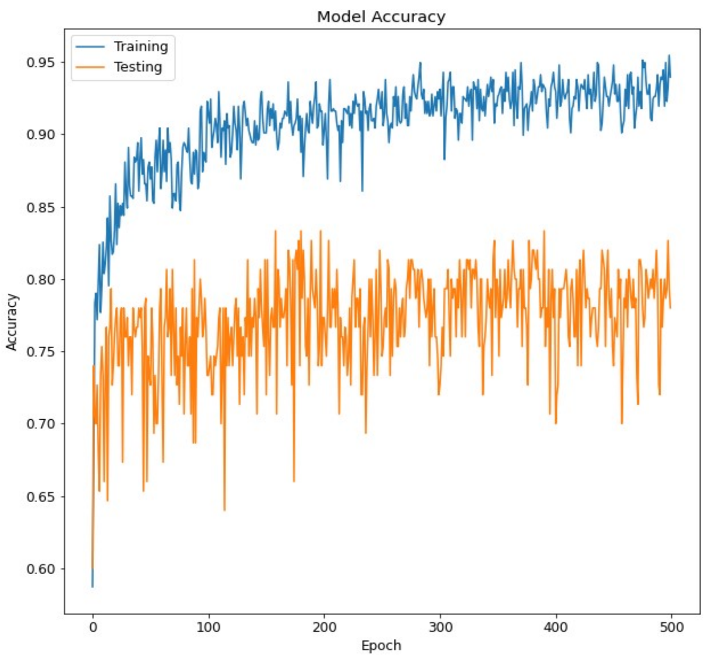

Figure 6.25 Accuracy Model of Xception Model using Chest X-Ray Images

B. Inception Model

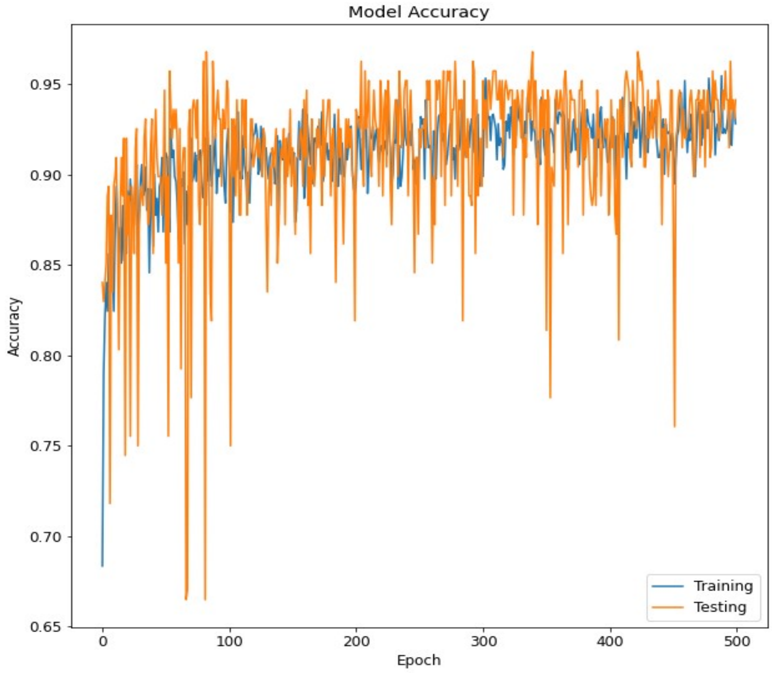

Figure 6.26 Accuracy Model of Inception Model using Chest X-Ray Images

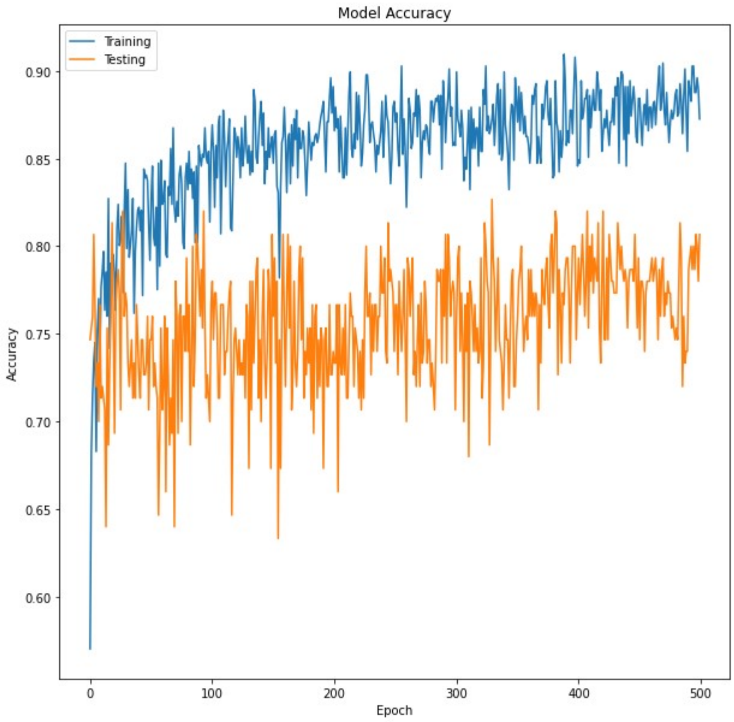

Figure 6.27 Accuracy Model of Inception Model using CT Scan

C. VGG Model

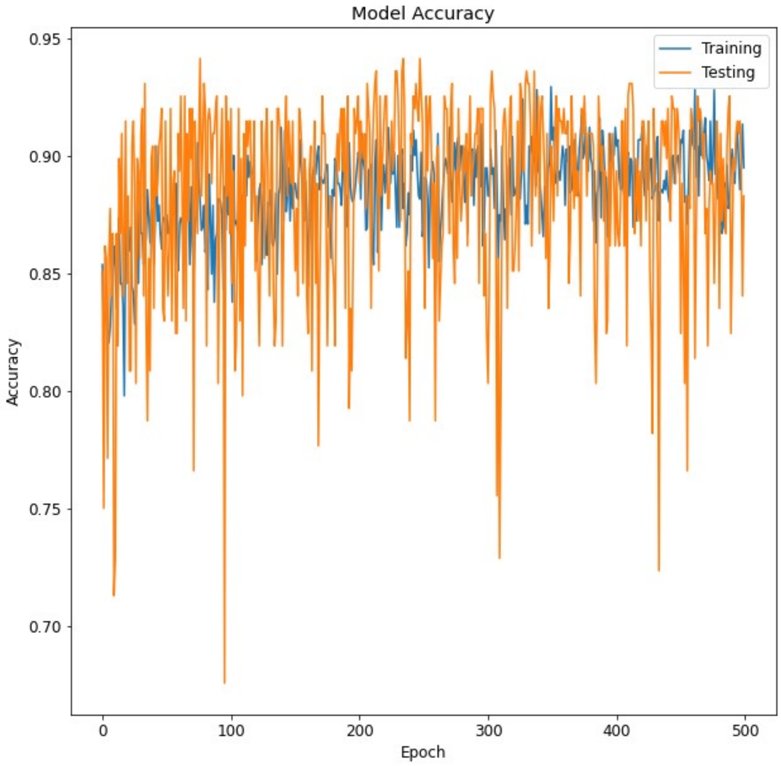

Figure 6.28 Accuracy Model of VGG Model using Chest X-Ray Images

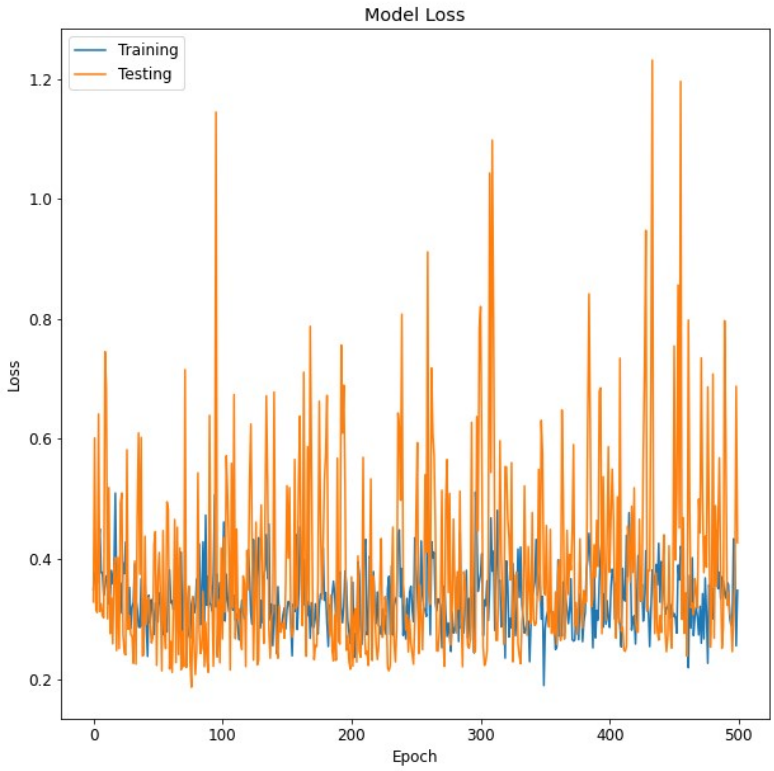

Figure 6.29 Accuracy Model of VGG Model using CT Scan

D. RESNET Model

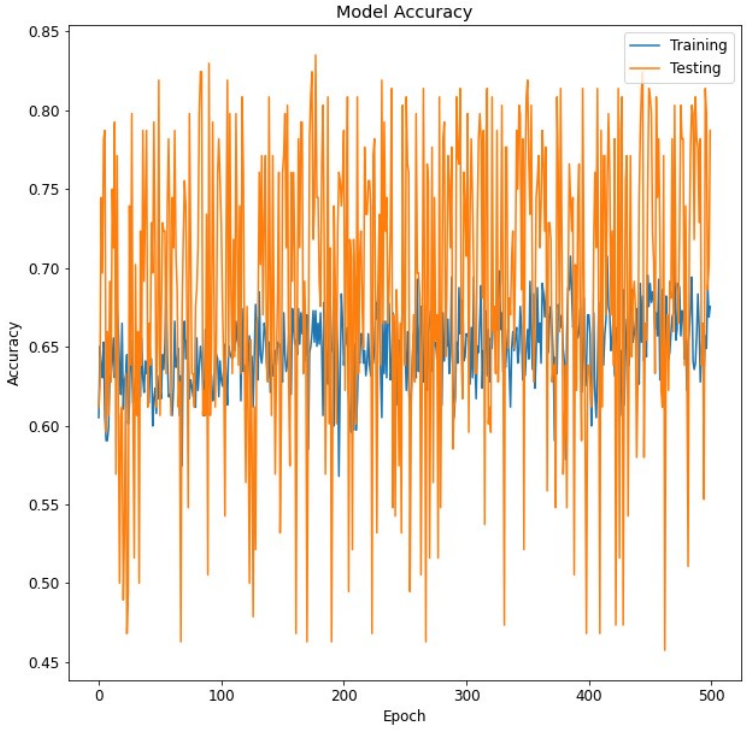

Figure 6.30 Accuracy Model of RESNET Model using Chest X-Ray Images

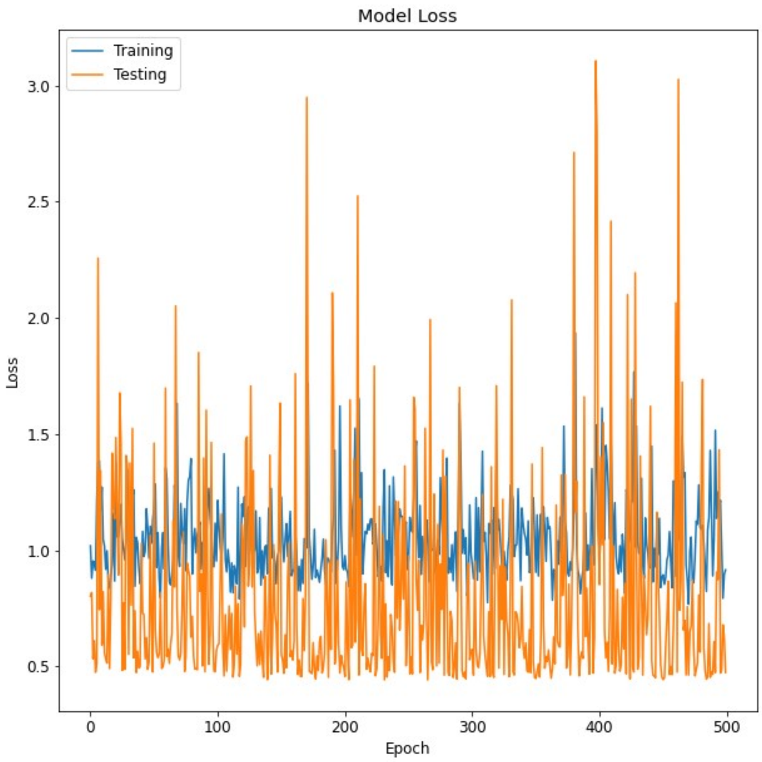

Figure 6.31 Accuracy Model of RESNET Model using CT Scans

6.3. PERFORMANCE ANALYSIS

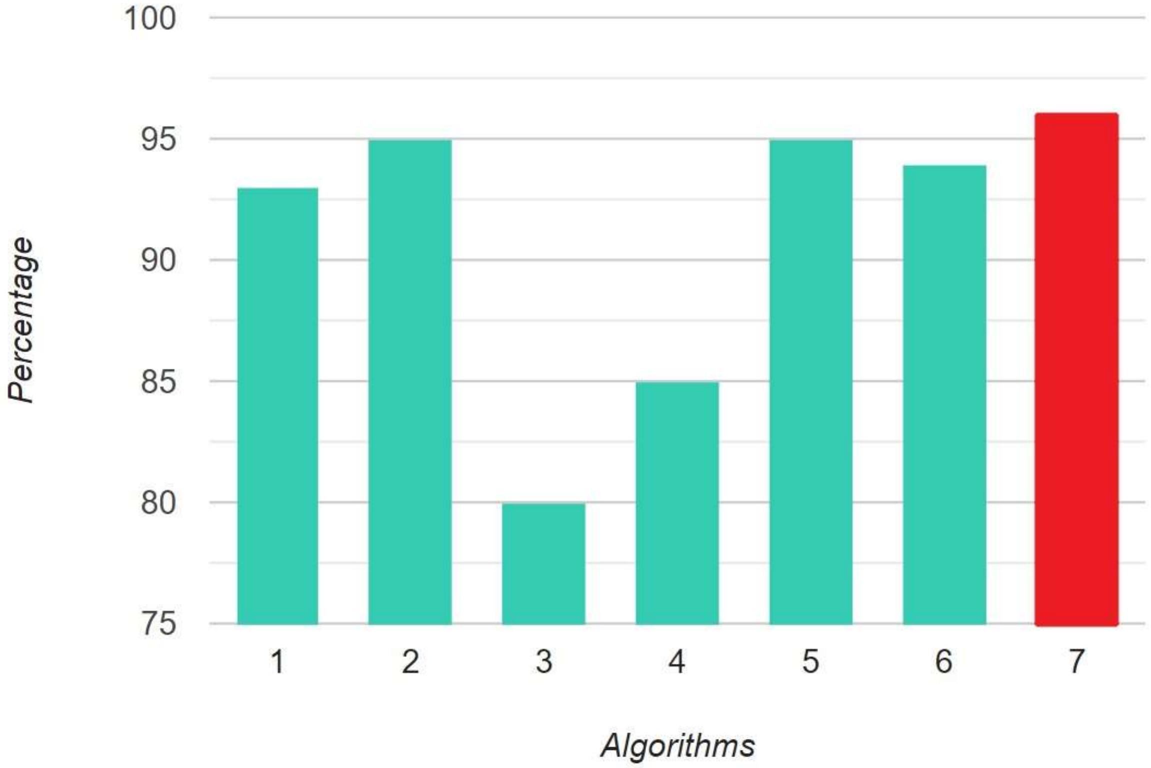

Figure 6.32 Performance Analysis

There are numerous image processing algorithms available for various applications, and the best algorithm to use depends on the specific task at hand.Here the numbers on the y-axis denote the percentage of the accuracy of the algorithms and the x-axis denotes the previous algorithms used in other systems and some in the literature survey.Most ot the previous algorithms on the x-axis as 1,2,3,4,5,6 are based on the CNN algorithm and their accuracy of each of the algorithms were 93,95,80,85,95,94 percentage accordingly. The performance of the previous algorithms is still impressive but there is a significant difference between is that this proposed system has a little increased accuracy than the previous algorithms with an accuracy of 96 percentage which is denoted by the number 7 on the x-axis.The analysis showed an output percentage which was more accurate than the previous observations .This is achieved by the inception algorithm model.

CHAPTER 7
CONCLUSION AND FUTURE WORK

Reliable recognition of infections in the lung is a key step in the diagnosis of Pneumonia disease. X-ray imaging examination of Chest is usually performed by trained human examiners or doctors, making the process time-consuming and hard to standardize. This research proposed and developed a Pneumonia detection model using the Deep Convolutional Neural Network and Pneumonia Chest X-ray dataset. This data was collected from the various patients and clinically examined and categorized by human examiners. The performance of the proposed model used, evaluated thus using different metrics such as Classification accuracy, Sensitivity, Specificity and the F1 score. The Classification accuracy of the proposed model achieved the average accuracy of 98.55% percentage in unseen chest X-ray images.

One potential future direction is to integrate the pneumonia detection algorithm with EHRs to enable real-time detection of pneumonia cases. This could allow healthcare providers to quickly identify patients who may be at risk of developing pneumonia and provide appropriate treatment. Another potential future direction is to investigate the use of multimodal imaging data (e.g., combining chest X-ray and CT images) to improve the accuracy of pneumonia detection. This could involve the use of advanced deep learning architectures that can fuse information from multiple modalities. Another potential future scope is to investigate the use of multimodal imaging data (e.g., combining chest X-ray and CT images) to improve the accuracy of pneumonia detection. This could involve the use of advanced deep learning architectures that can fuse information from multiple modalities.

REFERENCES

1. Aniket Gaikwad et.al.(2016)” Lung cancer detection using digital Image processing On CT scan Images”Issue: 04 | Apr-2016.
2. Arnold F. M et.al2020)‘‘Older adults hospitalized for pneumonia in the United States: Incidence, epidemiology, and outcomes,’’ J. Amer. Geriatrics Soc., vol. 68, no. 5, pp. 1007–1014, May 2020, doi: 10.1111/jgs.16327.
3. Attaway H et.al,(2021)‘‘Severe COVID-19 pneumonia: Pathogenesis and clinical management,’’ BMJ, vol. 372, p. n436, Mar. 2021, doi: 10.1136/bmj.n436.
4. Chan H et.al2020) ‘‘Computer-aided diagnosis in the era of deep learning,’’ Med. Phys., vol. 47, no. 5, pp. e218–e227, May 2020, doi: 10.1002/mp.13764.
5. Cherney K, (2022) “MRI vs. X-Ray: What you need to know Healthline Accessed”,: https://www.healthline. com/health/mri-vs-xray , 2022.
6. Hao Ren et.al(2021)“Interpretable Pneumonia Detection by Combining Deep Learning and Explainable Models With Multisource Data”,DOI: 10.1109/ACCESS.2021.3090215, 2021.
7. Jaiswal et.al (2021). Pneumonia Detection Using Xception, Inception, VGG and ResNet Model. Journal of Physics: Conference Series, 1844(1), 012080. doi: 10.1088/1742-6596/1844/1/012080
8. Jianpeng Zhang et.al2021)”Viral Pneumonia Screening on Chest X-Rays Using Confidence-Aware Anomaly Detection” ,PMID: 33245693 DOI: 10.1109/TMI.2020.3040950,2021.
9. Kaimal K. N et.al (2020). Pneumonia Detection from Chest X-rays using

Deep Learning Techniques. In 2020 International Conference on Artificial Intelligence and Signal Processing (AISP) (pp. 1-5). IEEE. doi: 10.1109/AISP49841.2020.9382318

10. Liu et.al (2020). Pneumonia Detection from Chest X-rays Using Convolutional Neural Networks with Transfer Learning. Journal of Medical Systems, 44(8), 152. doi: 10.1007/s10916-020-01569-y
11. Miotto R ,et.al(2018) "Deep learning for healthcare: Review, opportunities and challenges"PMID: 28481991 PMCID: PMC6455466 DOI: 10.1093/bib/bbx044,February 2018.
12. Mohammed Yseliani et.al (2022) "Pneumonia Detection Poroposing a Hybrid Deep Convolutional Neural Network" DO - 10.1109/ACCESS.2022.3182498,2022.
13. Nahiduzzaman M D et.al (2021) "A Novel Method for Multivariant Pneumonia Classification Based on Hybrid CNN-PCA"October 2021. DOI: 10.1109/ACCESS.2021.3123782.
14. Ozturk S, et.al (2020). Classification of COVID-19 and Pneumonia in Chest X-ray Images Using Convolutional Neural Networks and Transfer Learning. Applied Soft Computing, 93, 106476. doi: 10.1016/j.asoc.2020.106476
15. Rajput et.al (2021). Comparative Analysis of Pneumonia Detection Using Deep Learning Techniques. International Journal of Advanced Science and Technology, 30(5), 4965-4973.
16. Suguna S et.al,(2017) "A Survey On Prediction of Heart Diseases Using Big Data Algorithms", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 6, Issue 3, March 2017,ISSN:22781323.
17. Sun Y,et.al(2019) ''Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review,'' Sci. Rep.,

vol. 9, no. 1, p. 7600, Dec. 2019, doi: 10.1038/ s41598-019-44145.

18. Torres et.al(2021) ‘‘Pneumonia,’’ Nature Rev. Disease Primers, vol. 7, no. 1, p. 25, Dec. 2021, doi: 10.1038/s41572-021-00259-0.
19. Vinitha S et.al (2018) Disease Prediction Using Machine Learning Over Big Data, Computer Science & Engineering: An International Journal (CSEIJ), Vol.8, No.1, February 2018.
20. Wazif Khan, et.al (2021),“Intelligent Pneumonia Identification From Chest X-Rays: A Systematic Literature Review”,DOI: 10.1007/978-981-13-8798-2_1, 2021.