Thoracic Complication Detection on Chest X-rays using Deep Learning

by

Rizwanul Hoque Ratul
17101079
Farah Anjum Husain
17101166
Tajmim Hossain Purnata
17101146
Rifat Alam Pomil
17101284
Shaima Khandoker
17101003

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
June 2021

Declaration

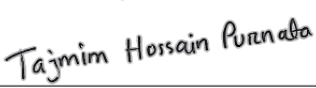
It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

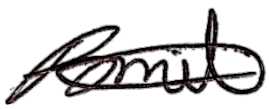
Student's Full Name & Signature:

Rizwanul Hoque Ratul
17101079

Farah Anjum Husain
17101166



Tajmim Hossain Purnata
17101146



Rifat Alam Pomil
17101284

Shaima Khandoker
17101003

Approval

The thesis titled "Thoracic Complication Detection on Chest X-rays using Deep Learning" submitted by

1. Rizwanul Hoque Ratul (17101079)
2. Farah Anjum Husain (17101166)
3. Tajmim Hossain Purnata (17101146)
4. Rifat Alam Pomil (17101284)
5. Shaima Khandoker (17101003)

Of Spring, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineer on June 6, 2021.

Examining Committee:

Supervisor:
(Member)

Zavid Parvez

Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.06.10 19:43:57 +10'00'

Mohammad Zavid Parvez, PhD
Assistant Professor
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

We hereby state that all the work done in preparing this thesis, beginning from the research to implementation of the proposed system and models is our own. Any inspiration from sources, datasets used and all other external resources that we used have been acknowledged. We have also refrained from using any unethical means in obtaining any of the resources used. We hereby testify that this paper has not been previously submitted in whole or part by anyone or any team to any other universities or institutes to complete a degree.

Abstract

Respiratory or thoracic diseases are one of the leading causes of death worldwide. Many of these diseases can be treated and prevented with proper diagnosis and early care. The diagnosis of thoracic diseases is mainly done with the use of chest X-rays and other chest imaging techniques and making sense of these require expert radiologists, who aren't always accessible. Especially in underdeveloped countries, a lot of patients with thoracic diseases are left to die due to a lack of proper diagnosis. With the advent of Artificial Intelligence, there have been many initiatives trying to tackle these medical diagnosis problems through the use of machine learning techniques and pattern recognition. Using Deep Neural Networks, patterns corresponding to different thoracic diseases can be easily detected from chest X-rays. In this paper, we have proposed just such a model that can identify the presence of a disease from a range of 14 different thoracic diseases using a Dense Convolutional Neural Network. Our dense convolutional neural network model takes advantage of the sheer amount of data that is now publicly available from chest X-ray datasets. This novel approach works in 2 stages, first training on images from disease-ridden patients alone, and then training the entire network on the whole dataset which includes X-rays from both healthy and unhealthy patients. This allows the model to make better predictions in detecting the presence of diseases as well as the absence. Given a chest X-ray alone, our model can give accurate predictions, with an AUROC mean score of 82.9% competing with the existing state-of-the-art models in this field.

Keywords: Deep learning; Chest X-ray; Convolutional Neural Networks; Thoracic diseases; DenseNet;

Acknowledgement

Firstly, we thank the Almighty for allowing us to move forward with this research even through all the difficulties we have faced during the pandemic this year and the previous. Without divine blessings, we would not be here. We also want to thank our supervisor, for guiding us in this journey, correcting our mistakes, and giving us the necessary feedback to improve our work ethic. He has given us valuable advice and made time for us out of his busy schedule. Furthermore, we want to thank our peers for helping us out whenever possible and giving us support. And finally, we thank our parents for their tremendous love.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AI Artificial Intelligence

$CADx$ Computer-Aided Diagnosis

CNN Convolutional Neural Network

CXR Chest X-ray

NIH National Institutes of Health

$PACS$ Picture Archiving and Communications System

WHO World Health Organization

Chapter 1

Introduction

The thoracic cavity or the overall chest area is home to some of the most crucial organs of the human body. It contains the lungs which supply us with oxygen and delivers oxygen to the whole body and most importantly the brain; without it, we cannot survive. The thoracic region also contains the heart, which is another crucial organ that carries oxygen provided by the lungs through the whole body using the circulation of blood. The survival and well-being of humans, therefore, is heavily dependent on the functions occurring in the thoracic region. If either the heart or lung has any problems, no matter how minute it may be, it can cause severe problems in your health. This region, especially the lungs, is also most susceptible to infections from external factors as it is constantly in direct contact with toxins, chemicals, and harmful particles from the environment through the air that we breathe in [37]. It is estimated that each year, about 4 million people die due to respiratory diseases and this number is increasing every year [40]. Respiratory diseases and thoracic diseases in general account for five of the most common reasons for death and chronic illness worldwide [37]. Many of these deaths are avoidable, as thoracic diseases can be widely prevented or treated.

The first step to the treatment of any disease is proper diagnosis. Respiratory or thoracic diseases such as Emphysema, Pneumonia, Infiltration, etc are usually diagnosed with the help of microbiological tests or different imaging techniques such as chest X-rays (CXRs) or CT scans[31]. Even diseases of the heart like Cardiomegaly are initially diagnosed with the help of CXRs. But these images can be difficult to interpret without the presence of an expert and this often leads to misdiagnosis and therefore mistreatment of diseases, causing the condition to worsen. It is, therefore, crucial to properly interpret CXRs and correctly diagnose these diseases. In this paper, therefore, we propose a system to automate this process and to help doctors better diagnose thoracic diseases with the help of artificial intelligence (AI).

1.1 Problem Statement

In this era of computers, machines, and a rapidly evolving world system, we are all heavily dependent on technology. From the simplest tasks to the heftiest ones, we can depend on machines to do the work for us. Technological advancements not only make our lives easier but also allow us to perform more complicated and

specialized tasks efficiently. An example of the scenario in question can be Radiology.

Diagnostic Radiology is the branch of medical science that deals with the diagnosis of diseases by the use of medical imaging techniques such as X-rays, CT Scans, MRIs, etc [36]. Most doctors for many years now have been using radiology to diagnose diseases and help treat them effectively [38]. This advancement in medical technology has been revolutionary since we can now detect changes or complications in the internal organs from outside the body without the need for surgery or even incisions. This is especially helpful when it comes to the diagnosis of thoracic diseases. These diseases are primarily diagnosed with the help of radiological tools. Currently, the imaging is done using machines [19], however, without the diagnosis of a specialist, the radiological reports have little value[38].

Diagnosis of respiratory and thoracic diseases from CXRs require expert Radiologists, and if a doctor, for instance, is not in their usual mental state or is tired from an entire day's work, they might fail to detect a small but crucial detail in the report of a patient[1]. Whether it be due to errors in diagnosis or some systemic fault in the machines used to generate these images to help test the patients, medical errors are very common and has been named as the third leading cause of death in the United States[11]. Especially in the case of thoracic diseases such as Pneumonia, Emphysema, etc. the identifying features of the diseases may be vague and can even overlap with other diseases[35]. So chances of misdiagnosis are further increased, as the human eye is less susceptible to such minute details.

Furthermore, trained Radiologists who can properly interpret CXRs might not be available at all times and in all places [1]. According to a report by the World Health Organization (WHO), as much as two-thirds of the world's population lacks access to advanced radiological diagnostics and specialists[35]. In lower-income countries, it gets even harder to find trained Radiologists to look over a CXR and diagnose diseases. Even when the radiologists are available, an appointment can be expensive. Thus many treatable thoracic diseases are left to worsen the condition of the patient, and patients suffer due to a lack of proper diagnosis. All these factors can equate to a higher mortality rate in the fate of these often very treatable diseases. We, therefore, propose a system to automate the whole process that will aid the doctor in diagnosing and analyzing CXRs more efficiently.

1.2 Research Objective

The main objective of this research is to build and analyze the performance of an automatic radiological diagnosis tool to assist doctors. Our model implementation is based on the concept of Deep Learning and will allow us to detect precisely the appearance of signs or anomalies in CXRs to be able to identify different cases of thoracic diseases from a range of 14 different pathologies [13].

The most popular approach to analyzing medical images has been through image segmentation using Convolutional Neural Networks (CNN). For the default CNN model, the simple yet effective Network-In-Network model is used because the model

is small in size, fast to train, and achieves similar or better performance than previous models[29]. The use of CNNs to diagnose the disease of a patient is expected to be at an accuracy level of a professional Radiologist, if not better[23].

Our main objectives therefore are :

- To develop a working system that can take in CXRs as input and label the presence of 14 different thoracic diseases (if any) accurately.
- Equate the probability of the disease most likely to be present and report disease diagnosis accordingly.
- Develop a system that can help distribute the load of a Radiologist and be implemented in the medical field to help diagnose diseases and reduce mortality rates.

1.3 Research Methodology

CXRs alone can be used to diagnose a plethora of thoracic diseases such as Mass, Cardiomegaly, Effusion, Atelectasis, Emphysema, Infiltration, Pneumonia, Pneumothorax, Consolidation, Nodule, Fibrosis, Edema, Pleural Thickening, Hernia, etc.[13]. With the recent developments in data reservoirs, there are now large amounts of such CXRs available for public use. This makes this the perfect problem to be solved using neural networks and can be used to train and develop deep learning models that can help diagnose thoracic diseases and automate the entire process. With the availability of such data on the internet, it is now easier than ever before, to accurately train neural networks so that they can detect patterns in a wide range of cases. One such publicly available dataset is the National Institute of Health (NIH) ChestX-ray 14 dataset [39][28], which gives us access to 112,120 labeled chest X-rays[15].

In our model, we propose a novel multi-stage optimization approach to diagnosing thoracic diseases from CXRs by training a CNN on the NIH dataset so that it can identify the important features that equate to the presence of thoracic diseases. In the first stage, the model is trained on CXRs that have been labeled with the presence of thoracic diseases, i.e. CXRs of non-healthy patients. And in the second stage, we train the model on the entire dataset including both sick and healthy patients' CXRs. The neural network, once trained, should be able to correctly label the presence of a disease given just the CXR of the patient. Our model implementation is based on the DenseNet-121 structure[16] and inspired from the other existing models in this field.Using transfer learning, we have combined different features and have designed a model that works in two stages and several sub-stages or phases to give better results.

1.4 Thesis Outline

Our research aims to develop a working system that can detect the presence of thoracic diseases from a range of 14 different thoracic diseases which include Mass, Cardiomegaly, Effusion, Atelectasis, Emphysema, Infiltration, Pneumonia, Pneumothorax, Consolidation, Nodule, Fibrosis, Edema, Pleural Thickening and Hernia from a CXR of a patient. We have used the ChestX-ray14 dataset from NIH to train our model in two different stages to properly predict the presence and/or absence of these 14 diseases. The overall paper explores these steps that were followed to develop this working model.

Chapter 1 outlines our inspiration for choosing this topic and discusses the problems that are to be addressed with our research.

Chapter 2 deals with relevant work in this field by others. We look deeper into the various methods and approaches to detecting thoracic diseases and explore the potentials of these kinds of systems in the medical field.

Chapter 3 describes the dataset and its nuances and incompetencies. We go in depth into how the dataset was collected and how we processed it for use in our model.

Chapter 4 proposes the model structure that we have used to implement our research. We go into detail about how each aspect of the model works together to predict results, starting from analyzing the data to how the two stages work together to produce better accuracy.

Chapter 5 analyzes how our model performs and compares the results with the state-of-the-art models in this field. We also discuss the scope of our model, address our limitations and discuss future work.

Chapter 6 concludes our proposal and talks about the future potential for this research.

Chapter 2

Literature Review

With the help of CXRs, we can now easily observe the internal condition of the lungs and bones that exist in the chest area. They allow the detection of anomalies or diseases of the airways, blood vessels, bones, heart, lungs, and whether there is a presence of fluids in the lungs[25]. These features of abnormalities can be indicative of diseases like Pneumothorax, Consolidation, Atelectasis, Cardiomegaly, Effusion, Pleural Thickening, Infiltration, Mass, Nodule, Pneumonia, Edema, Emphysema, Fibrosis, and Hernia[36]. Just recognizing these slight changes in an X-ray can be a crucial point in the diagnosis of a disease.

Computer-Aided Diagnosis (CADx) systems have already been around for a while, and are clinically approved, and proven to decrease false detection rates[2]. However, unlike CAD systems, which just highlight the anomalies, deep learning can extract useful features which are beyond the limit of radiology detection[27]. Furthermore, the system should increase in accuracy as more and more data is fed into it, and it evolves with time.

Perhaps the most prominent work in this area has been by Wang et al(2017) who have become the benchmark for all other following approaches to working on diagnosing diseases from CXRs[15]. They were the first to work with such huge amounts of data in this field, release the database (that paved the way for future work by multiple others), and also popularize the use of pretraining models on ImageNet for better results[4]. They showcased that it was possible to detect diseases and spatially locate them using a "unified, weakly supervised, multi-label image classification and disease localization framework". As Wang et al[15] is part of the NIH, they have used their institute's CXR image database to develop this labeled dataset with 14 different diseases. Using text-mining techniques each of the images is labeled with which diseases they correspond to. A certain number of images they have provided also come with hand-labeled bounding boxes that can help with evaluating disease localization performance.

Ever since large data sets like the ChestX-ray 14[13][14][15] and MIMIC-CXR[26] have been available there has been a large number of research papers and projects attempting to computerize the diagnosis process. Just focused on CXRs alone, there are a variety of approaches to using Convolutional Neural Networks to label and diagnose these diseases. One such approach is by Mahmud Monshi et al (2019), where

the authors propose a deep convolutional neural network architecture, based on ResNet50, to detect the presence (or absence) of twelve thoracic diseases[27]. Their model was tested on the MIMIC-CXR dataset[26], using eight epochs and a subset of the multiview CXRs that are available, to detect 12 common thoracic diseases which include: pneumonia, lung lesion, enlarged cardiomediastinum, cardiomegaly, airspace opacity, edema, consolidation, atelectasis, pneumothorax, pleural effusion, etc.

Their system works in 3 stages, with a total of 12 sub-tasks each considering the presence or absence of one disease. In the first stage, they train the ResNet50 model, using a transfer learning approach to train faster with a model that is already trained to recognize 1000 categories of things in ImageNet. The second stage calls the fit-one-cycle method, to observe the model's performance and increase accuracy. The final stage trains the model again for 4 epochs, using an optimal learning rate finder. The experiments are performed using 10% of the MIMIC-CXR dataset, using all available frontal and lateral views of the chest X-rays using small image sizes of 224 by 224 pixels and resulted in an average AUC score of 0.7999[27].

Another similar study has been done by Hongyu Wang and Yong Xia (2020), in which they have used the ChestX-ray14 dataset along with an additional 180,000 images from the Prostate, Lung, Colorectal, and Ovarian or PLCO dataset from NIH[5] for more training data[33]. Their neural network model, named ThoraxNet, consists of two branches working together to diagnose the 14 diseases. The Classification Branch is used for Label Prediction and is just the ResNet-152 model adapted to their problem by removing the softmax layer and replacing the last fully connected layer with a 14 neurons layer, each of which uses the sigmoid activation function. The other branch is used to detect abnormalities and is called the Attention Branch. This branch is used to exploit the correlation between class labels and the regions of pathological abnormalities, using the output of the residual module in ResNet 152 as the input layer, by analyzing the learned feature maps. The diagnosis is obtained by averaging and binarizing the outputs of both branches and achieves an AUC score of 0.7876.[33].

A common limitation of most approaches is that they focus only on the X-ray image, and ignore any other radiological reports which can often be crucial in the diagnosis process. TieNet, developed by Xiaosong Wang et al. (2018) has also considered this and proposed a system that stimulates the real-world reporting process by outputting disease classification and generating a preliminary report spontaneously[19]. Their approach to this problem involves the use of a multi-level attention model[14] integrated into an end-to-end trainable Convolutional Neural Network - Recurrent Neural Network architecture that highlights the meaningful text words and relevant image areas. The text embedding learned from the model is incorporated into the rest of the model as a priori knowledge.

The Recurrent Neural Network (RNN)[3] used in this model is based on a visual image spatial attention model for image captioning. This RNN accepts the current word at each time step as input, as well as the soft-weighted visual features. It uses attention to combine the most salient portions of the RNN hidden states. It

re-uses the attention mechanism except that it performs a max-over-r operation, producing a sequence of saliency values for each word. These saliency values are used to weight and select the spatial attention maps generated at each time point. This system annotates the images from the Chest X-ray 14 dataset, hand-labeled images, and the OpenI[8] dataset. With global representations computed for both the image and report, these are combined to produce the final classification which is the simulation of a text report that a radiologist would write[19].

Another disadvantage most machine learning systems face with analyzing and diagnosing diseases is the lack of pixel-wise annotated data by radiologists or coarse bounding boxes in sufficient amounts. Chaochao Yan et al.(2018) in their paper, have therefore devised a weakly supervised deep learning system to aid radiologists with unlabeled data[20]. The proposed system classifies thoracic diseases merely by reading provided CXRs as well localizes the disease regions on the CXRs at pixel-level granularity. The system produces disease-specific results instead of treating all diseases as the same. In this paper, the publicly available Densenet-121[16] model has been used as a backbone network because it consists of four consecutive dense blocks. At first, a Squeeze-and-Excitation step is done which takes the advantage of the widely existing cross-channel dependency. In between two consecutive dense blocks of DenseNet[16], there is a convolution-pooling operator where the squeeze and excitation block is inserted. Once that step is done, the system moves onto Multi-map Layer and Max-Min Pooling. The Multi-map layer is required due to CXR having multiple disease labels which requires multi-class classification therefore each output feature map corresponds to a particular disease class. To sufficiently utilize the provided multi-class level, the Max-Min Pooling was introduced to aggregate information on feature maps for each disease class[20].

In the present world, perhaps the most substantial application of a Neural Network model that diagnoses CXRs would be the relation of Pneumonia to Covid-19. While the exact percentage of Covid-19 patients who get diagnosed with Pneumonia is still uncertain, often the cases that do may lead to more serious consequences[34]. With the disease being a global pandemic as declared by the World Health Organization [32], and numbers skyrocketing each day, an automated system to help doctors save their precious time and effort might prove to be crucial. Tanvir Mahmud et al. proposes a model called CovXNet, a deep convolutional neural-network-based architecture, which extracts diversified features from chest X-rays of confirmed COVID-19 patients[30]. Due to the significant similarities between viral/bacterial pneumonia and COVD-19 caused pneumonia, the system is initially trained using l1583 normal CXRs, 1493 non-COVID viral pneumonia CXRs and 2780 bacterial pneumonia CXRs collected in Guangzhou Medical Center, China[18]. Once the initial training is done, the model is then further specialized using a radiologist verified database containing 305 CXRs of different COVID-19 patients collected from Sylhet Medical College, Bangladesh. Different resolutions are used to improve accuracy and a stacking algorithm is also employed. The proposal uses a gradient-based discriminative localization to differentiate between abnormal regions of X-rays referring to the different types of pneumonia. While the system has produced an accuracy of 97.4%, it is to be noted that the system is mainly trained on non-COVID data, as the data is still evolving and changing due to the situation.

Unlike the other papers aforementioned, the CovXNet approach rather focuses on one specific disease instead of 13 others. The system as of yet is far from being used effectively in the medical world, as not enough compiled data is available in a way that can be used to teach the network. The current ChestX-ray 14 dataset may not be suitable for this as well, since the labels do not distinguish between Viral and Bacterial Pneumonia, which is a crucial factor in the diagnosis of Covid-19[34]. As the virus and disease are still evolving, the studies are simultaneously evolving and changing to account for the vast amounts of data being collected. As more and more data is available, many other systems will be better equipped to diagnose and deal with Covid-19 Pneumonia. Our aim however is slightly different, as we diagnose 13 other thoracic diseases as well as Pneumonia, but the system itself may be specialized to meet requirements.

While most problems in the medical imaging field are perfect candidates for implementation using neural networks and machine learning techniques, there are a lot of realistic constraints that become obstacles in these models being applied in the medical field. Since this is the question of human lives and one wrong prediction could cause a lot of damage, it is very important to consider all relevant aspects that deal with the accuracy of these predictions. One such important factor is that in a lot of the cases, the features presented in a CXR (or any medical imaging technique) are that the presence of a certain feature may indicate the presence of multiple diseases. It is not just a simple binary classification problem. They are in fact, multi-label classification problems. The paper by Yao et al incorporates this factor and uses LSTMs[9] to take advantage of multi-label dependencies[21]. They manage to get state-of-the-art results in predicting the 14 different diseases on the ChestX-ray14 dataset[13] [14] from NIH, and they do so without pre-training.

As can be seen, there are a lot of different approaches to solving the same problem. However, our work is inspired by that of Pranav Rajpurkar et al(2017)[13], in the CheXNet model which uses DenseNet [16] and batch normalization for optimization. The CheXNet model is the standard for most of the work in this path, as it has been widely acknowledged. The weight of this network is initialized from ImageNet's[4] pre-trained models and normalization is based on the mean and standard deviation of images from ImageNet. The system replaces the final fully-connected layer with a fully connected layer producing a 14-dimensional output, after which an elementwise sigmoid nonlinearity is applied.

Another recent proposal that is similar to our model is that of Hongyu Wang et al(2021)[35]. They have proposed a model that can detect 14 different thoracic diseases, just from a CXR. They have used a DenseNet-121 [16] architecture, pre-trained on ImageNet[4], as their base model, and have incorporated three different attention modules to focus on scale-wise, element-wise, and channel-wise learning. They train their model on the ChestX-ray14 dataset[13] [14] and yield the highest average per-class AUC of 0.826.

Different models aim to solve different issues in the diagnosis of thoracic diseases from CXRs, and in return have limitations of their own. One of the most important

problems of these systems is the disregard of patients' medical histories and current clinical records, which can often prove to be crucial in diagnosing a disease. Other problems include data augmentation and pixel normalization, lack of accurate localization of lesion areas utilizing the limited amount of bounding boxes, and not treating the diseases separately and thus ignoring that those lesion areas on chest X-rays are disease-specific. One other significant factor is that many images from the ChestX-ray14 dataset contain lesion areas of overlapping thoracic diseases. With such varied differences and problems to a singular solution, there are many strides left to be taken in improving computer-aided diagnosing systems for Chest X-rays and thoracic diseases. A more holistic approach would be necessary to make any of these systems applicable in the real medical world.

Chapter 3

Dataset Description and Pre-Processing

3.1 Data Collection

This is the age of data, and with such huge amounts of data publicly available, new doors of data science and research have opened up and every day new and improved ways of utilizing this data are being discovered. The availability of such huge data makes it perfect to be used in deep leaning methods.In our approach, we have chosen the ChestX-ray14 dataset [39] [28], or the "NIH Chest X-ray Dataset", which is publicly available on Kaggle, one of the largest data science communities that provide powerful and effective tools and resources for public use[39].

Before this release, the largest available dataset of CXRs was OpenI with only 4143 images available. Wang et al. [15] paved the way for hundreds of others to use this data and develop new ways to detect thoracic diseases. This dataset was first released by to utilize the huge data reserves of CXRs that NIH holds in their Picture Archiving and Communications System(PACS). One of the members was a fellow at NIH and another one was an employee and hence they had access to huge amounts of CXRs as well as their corresponding radiological reports from doctors.

The NIH PACS is huge and contains reserves of many different categories of reports. To create the database, first of all, a list of 8 pathological keywords that were most commonly found in thoracic disease diagnosis from radiologist reports was established with help from the radiologists at NIH. Then these 8 keywords were used to text mine the relevant radiological reports from the PACS reserve using Natural Language Processing (NLP). They also used NLP to remove any negation and uncertainty. The NIH has not released these radiological reports to the public and hence this dataset is our only gateway to using NIH's huge data reserve. The labels are claimed to be more than 90% accurate[28]. Wang et al[28] further demonstrated the application of this dataset with a disease localization framework as well. Figure 3.1 summarizes their methods. This then inspired many more researchers to attempt to solve this problem in their way.We are no different and it would have been nearly impossible for us to gather such huge amounts of data ourselves. This is why have chosen to utilize this public reserve and use it to train our neural network.

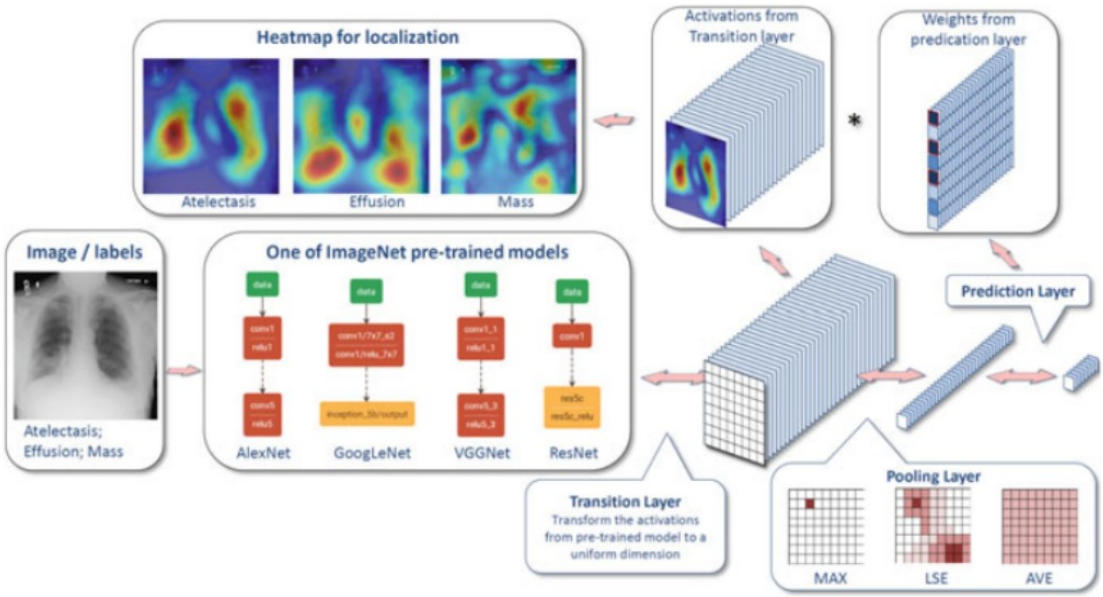


Figure 3.1: The deep convolutional neural network used to show the localization of diseases from CXRs, developed by Wang et. al. [28].

3.2 Description of the Dataset

The dataset consists of 112,120 CXR images with disease labels from 30, 805 unique patients, collected during the span of 1992 to 2015. The 14 common thoracic diseases that are labeled here include Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural thickening, Cardiomegaly, Nodule, Mass, and Hernia[28].

Since the data from the ChestX-ray-14 dataset [39] [28] is labeled with the use of NLP techniques, there are still some discrepancies in the data, as opposed to the accuracy of radiologist-labeled data[15]. With such a large dataset it is extremely hard to manually label all this data, especially since it involves the expertise of a Radiologist. It is expected, therefore, that there will be the presence of certain errors. The data is also rather imbalanced, as shown in Figure 3.2; for instance there are about 19894 X-rays for Infiltration whereas only 227 for Hernia. So the model is better trained on detecting Infiltration. Furthermore, out of the 112,120 images, more than 60, 000 images have been labeled with "No Findings"; which means that these are images with no presence of any of the 14 detectable diseases.

CXR or any form of medical imaging contains certain areas or tell-tale signs that help the doctor diagnose the disease as noted in Figure 3.3. Wang et al [28] have also shown that these diseases can be spatially located using localization techniques. Using pattern recognition and deep learning to detect these areas rather than CADs is much more effective because of the minute details that can signify the presence of a disease. Our system also attempts to recognize these spatial patterns to classify the diseases.

Furthermore, often some of these thoracic diseases have connections to each other[28].

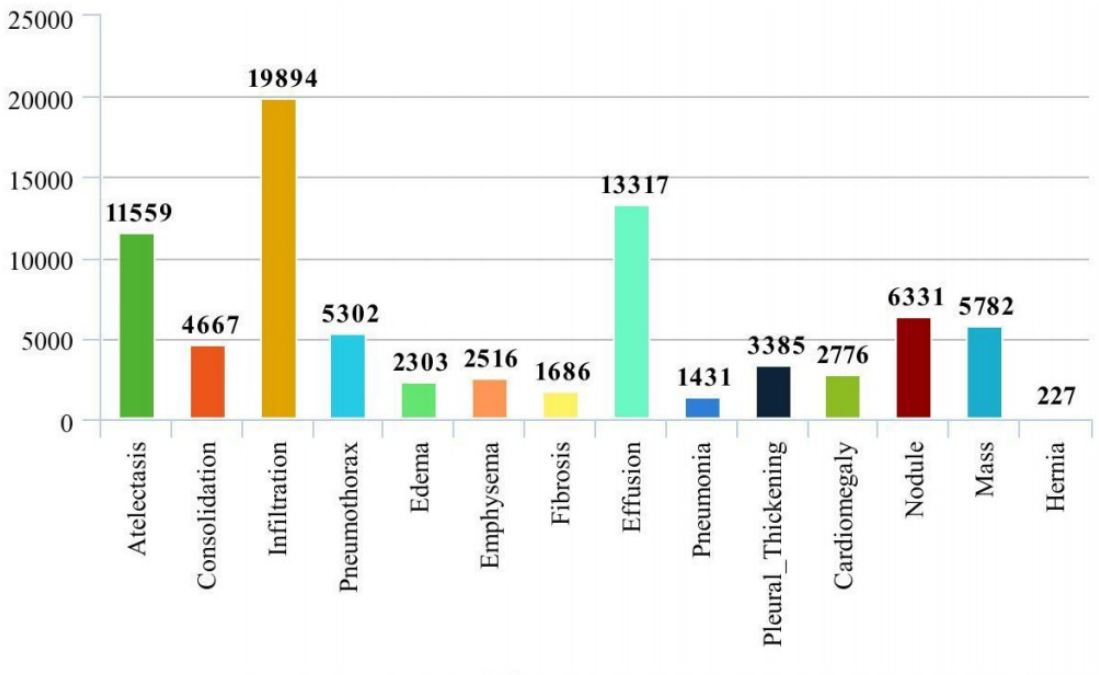


Figure 3.2: Disease distribution in the dataset. Evidently, there is inconsistency in the data; while Infiltration has 19894 occurrences, Hernia has just 227.

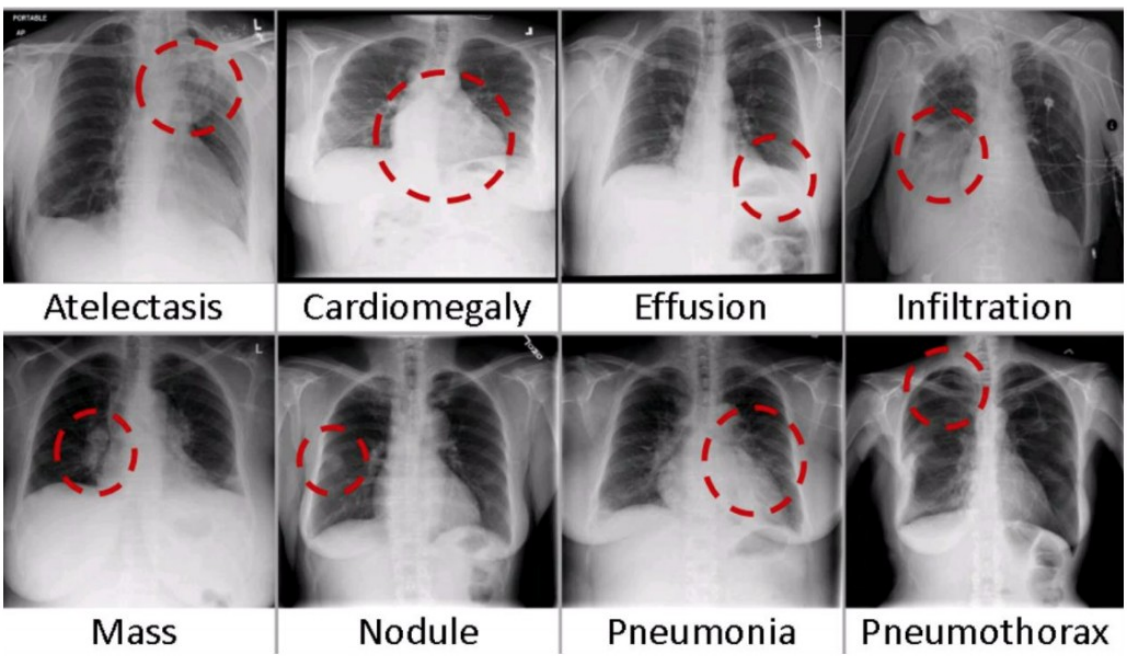


Figure 3.3: CXRs of patients with different diseases. Circled areas showing region-specific features correlated to the disease[15].

Figure 3.4: Proportional representation of disease co-occurrence statistics in the dataset released and labeled by Wang et al.[28].

This has been established by radiologists and it is also obvious from the PACS reserve. Thus in a lot of the cases, multiple diseases will be present and their corresponding spatial locations that signify their presence can overlap as can be seen from Figure 3.4. This overlapping is coherent with the radiologist's insights.

Certain combinations of pathologies always seem to appear together or with all kinds of pathologies. These include Effusion, Atelectasis, Cardiomegaly with Effusion, Emphysema with Pneumothorax, Nodule with Infiltration, Edema with Infiltration, Fibrosis with Infiltration, and Pneumonia with Infiltration. Multiple pathologies are therefore much more common in most of these CXRs as shown in Figure 3.5.

3.3 Data Pre-Processing

Since the dataset was developed for a classification task in the first place, it has already been through certain steps of necessary pre-processing and it isn't the same as raw data. As can be seen from Figure 3.6, the data is pretty organized and the image distribution has already been pre-determined with labels and corresponding information. We use the image distribution file developed by Rajpurkar et al.[13] The image distribution file is then merged with the NIH label set, which correlates the information with the images.

The original images with dimensions of about 2500x2800 are resized to 224x224, for ease of processing. They are resized so they can be easily used in the DenseNet-121 structure[16]. They are then further processed with the help of Dense Net-121 where horizontal flipping, standard normalization and a 15 degree rotation was performed and zoomed by 0.15 times to introduce more versatility into the data. This allows the system to better recognize the image even with different alterations or rotations.

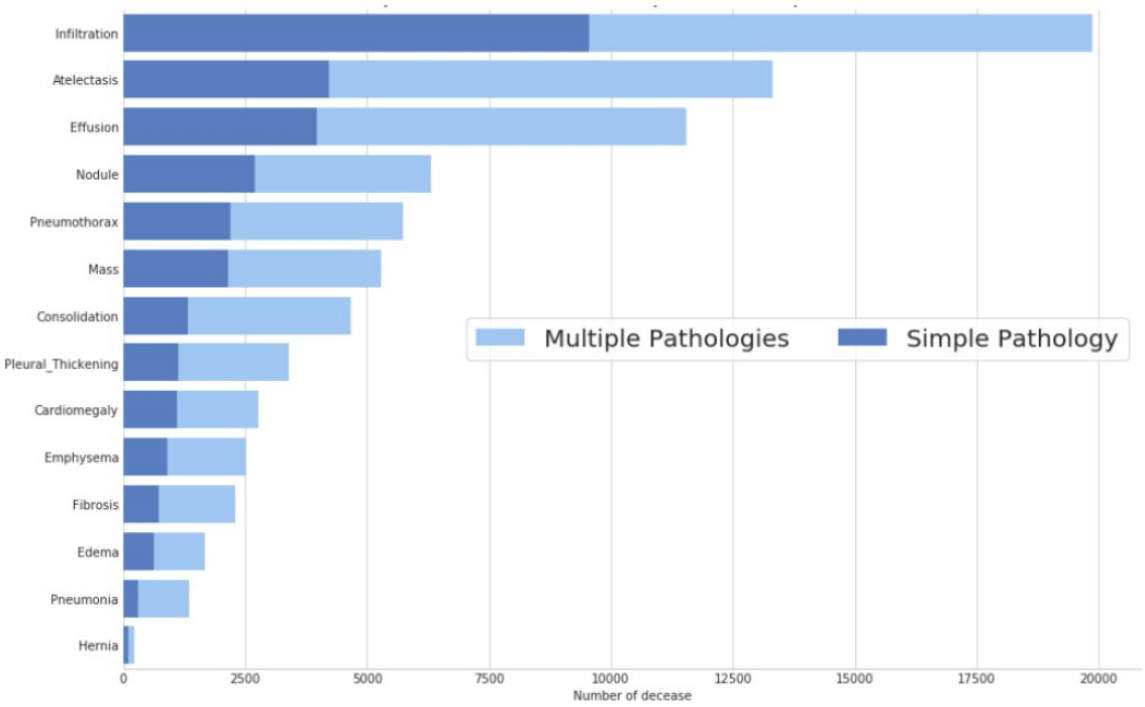


Figure 3.5: Occurrence of pathologies alone vs together.

The images and labels are fed separately into the model, with the labels containing the information of all the patients including gender, age, presence of disease (denoted with a 0 (absent) or 1 (present)), and the path to each corresponding CXR.

View_Position	Original_Image_Width	Original_Image_Height	Original_Image_Pixel_Spacing_X	Original_Image_Pixel_Spacing_Y	dfd	Atelectasis	Consolidation	Infiltration	Pneumothorax	Edem
PA	2682.0	2749.0	0.143	0.143	NaN	0.0	0.0	0.0	0.0	0
PA	2894.0	2729.0	0.143	0.143	NaN	0.0	0.0	0.0	0.0	0
PA	2500.0	2048.0	0.168	0.168	NaN	0.0	0.0	0.0	0.0	0
PA	2500.0	2048.0	0.171	0.171	NaN	0.0	0.0	0.0	0.0	0
PA	2582.0	2991.0	0.143	0.143	NaN	0.0	0.0	0.0	0.0	0
...	...	...	...	...	...	...	...	...	...	...
AP	2500.0	2048.0	0.168	0.168	NaN	0.0	0.0	1.0	0.0	0
PA	2938.0	2991.0	0.143	0.143	NaN	0.0	0.0	0.0	0.0	0
AP	2500.0	2048.0	0.168	0.168	NaN	0.0	0.0	0.0	0.0	0
PA	2048.0	2500.0	0.171	0.171	NaN	0.0	0.0	0.0	0.0	0
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na

Figure 3.6: A snippet of the Dataset Label File.

3.4 Training, Validation, and Testing

There are cases of multiple CXRs from one patients in the dataset. We have ensured that there is no overlapping in the train-validation-test data. This means, no two CXRs from the same patient are used in more than one stage. Each set of data fed into these 3 phases are different and belong to different patients so there is no bias involved in pattern recognition.

During the training phase, we use an image size of 224x224 and a batch size of 16, during both stages. As can be seen from Table 3.1, for the training data-set, we use

32, 422 images from about 12, 961 patients, during the first stage. For the second stage, we use 77, 946 images from 27, 724 patients. Augmentation is done using horizontal flipping, and standard normalization methods are also used.However, we do not use vertical flipping.

For the validation phase, we use 1440 patients and about 2599 images in the first stage. Whereas for the second stage we use 8575 from 3078 patients.

Finally, for the testing phase, the batch size is 8192, and we use about 15,734 images from 1286 patients during the testing phase for the first stage and 25,595 images from 2797 patients in the second stage. The batch size is still kept to 16 for both stages.

Stage	Train	Validation	Test
1st stage	32,422	2,599	15,734
Unique Patient	12,961	1,440	1286
Batch size	16	16	8,192
2nd stage	77,946	8,575	25,595
Unique Patient	27,724	3,078	2,797
Batch Size	16	16	8,192

Table 3.1: Splitting of the data-set

Chapter 4

Proposed Model

4.1 Architecture

Our model works in two stages, with slightly different data input sets. After resizing the original CXR images, only the CXRs with some label of existing pathology is fed into a pre-trained DenseNet-121 model for Stage 1. We do not feed the CXRs with a "No Findings" label into the neural net, which accounts for about 60, 000 images. Once weights are initialized, we freeze all the weights from lower convolutional layers, and then we begin stage 2. In stage 2 we then feed the entire dataset including CXRs from both healthy and sick patients. Then we replace the final fully-connected layer with a fully connected layer of a 14-dimensional output and treat the DenseNet-121 as a fixed feature extractor. In the second stage, we fine-tune the weights from all the layers by backpropagation. Each training iteration optimizes the cross-entropy losses.

4.1.1 CNN Layers and Functions

Before we dive into the working system, here is a summary of the different existing techniques we have used to develop our model.

DenseNet-121

DenseNet121 consists of 121 layers of Densely Connected Convolutional Neural Networks. Each layer in the model of the DenseNet architecture obtains additional inputs from all preceding layers and propagates its feature maps to all other subsequent layers. Due to this concatenation, all of the layers are receiving "collective knowledge" from all the preceding layers. This means the network built will be thinner and more compact thus improving the computational efficiency and memory efficiency[16]. The classifier in DenseNet gives smoother decision boundaries even when training data is insufficient because it uses features from all existing complexity levels.

Convolutional Neural Network

CNN is a popular class of deep learning neural networks. CNNs can be thought of as a machine learning algorithm that mainly works with image classification problems. It takes in an input image, assigns relative importance (learnable weights and

biases) to various aspects/objects in the image, allowing them to differentiate one feature from the other. CNN works by extracting features from the images[12]. All CNNs consist of three main layers- the input layer which is a grayscale image, the Output layer which is a binary or multi-class label, and the hidden layers consisting of convolution layers, ReLU (rectified linear unit) layers, the pooling layers, and a fully connected Neural Network. Each of the layers of a CNN has multiple convolutional filters working and scanning the complete feature matrix and carrying out dimensionality reduction. This enables CNNs to be a very apt and fit network for image classifications and processing[12].

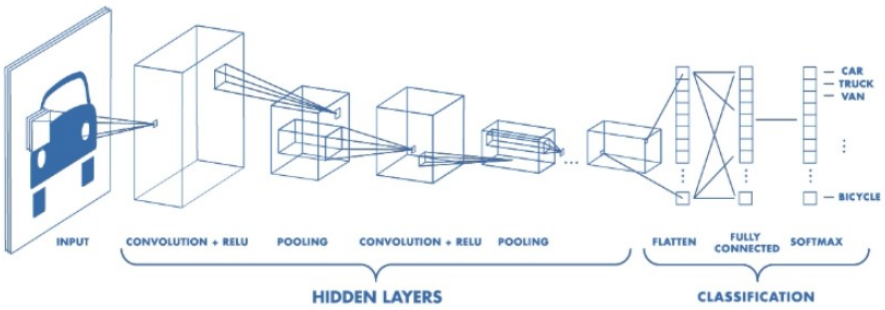


Figure 4.1: The layers of the Convolutional Neural Network Explained.[6]

We start by converting our images into gray-scale form because it makes it easier to process without losing important features that are necessary to make the algorithm scalable onto massive datasets.We can see the layers of a CNN in Figure 4.1

Convolution Layers

The convolution layer is made up of one or more filters (which is basically like a matrix) each with different weights that are used to extract features from the input image. In the case of convolutional layers, the neurons aren't all connected to each other, but only connected to a few significant input neurons so that there are less parameters to focus on [17]. This allows the entire network to go much deeper, using less parameters. The features for the various images are calculated based on their neighboring pixel values, while the filter slides over the input image. As the filter travels along the height and width of the image, the dot product between the filter matrix and the input image matrix is calculated at every spatial position. This is known as convolving. Each filter convolves or slides over all the pixels of the input image and extracts relevant features from it.

Adam Optimizer

Adaptive Moment Estimation or Adam is an optimization algorithm that we used to train the model with the CXRs[22]. Adam uses an exponentially decaying average of past squared gradients in order to produce an adaptive learning rate. It is similar to the momentum optimizer and the RMSProp optimizer as implied by its defining equations, (where $\odot$ denotes the Hadamard product and θ denotes Hadamard division):

$$
\begin{aligned}
m &= \beta_1 m - (1-\beta_1)\,\nabla_\theta J(\theta) \\
s &= \beta_2 s + (1-\beta_2)\,\nabla_\theta J(\theta) \odot \nabla_\theta J(\theta) \\
\hat{m} &= \tfrac{m}{1-\beta_1^t} \\
\hat{s} &= \tfrac{s}{1-\beta_2^t} \\
\theta &= \theta + \eta \hat{m} \oslash \sqrt{\hat{s}+\epsilon}
\end{aligned}
\tag{4.1}
$$

ReLU Function

ReLU or rectified linear unit is a form of the activation function that is used to increase the non-linearity of the network[24]. However, it does not affect the receptive fields of each of the convolution layers, while doing so. ReLU allows faster training of the data. This is introduced after each convolution layer to introduce non-linearity in the feature maps. It ignores negative values and only outputs the positive values or zero.

$$ReLU = max(0, value) \tag{4.2}$$

Max Pooling

The max pooling layer applies a non-linear down-sampling on the convolved feature often referred to as the activation map [7]. This is done so that the computational complexity is reduced, as we are dealing with large amounts of data that is linked to an image. This is performed by sliding a window over the image selecting the average, maximum or minimum values in the window depending on the task at hand. So that reduces further complexity of the network by minimizing the number of parameters.

Image Flattening

The neural network can work on a tabular structure so the output of the pooling needs to be converted to a compatible format so that it can be classified. The data is therefore converted into a one-dimensional array. Also often a dropout layer is added to prevent overfitting of the algorithm. Dropouts prevent overfitting by reducing the correlation between neurons.

Fully Connected Layer

This is the last layer, where the feature maps are flattened into a one-dimensional array and dense layers are connected to it i.e. every neuron in the current layer is connected to every other neuron in the next layer. A softmax classifier at the end is used to finally classify the image into the given number of classes[10].

So to summarize, the flattened feature maps obtained from DenseNet-121 will be inputted into the CNN in the form of a dataset. The CNN will carry out supervised learning by using an adequate number of convolutional and pooling layers. Finally, it will classify the feature maps into correct classes of disease.

4.1.2 Developing the Model

The task of diagnosing or detecting the presence of certain diseases from a CXR is first and foremost a binary classification task. This is to say, that in the chance of an actual implementation of the thoracic complication detection model, the resulting model, given a CXR, should accurately predict whether or not there is the presence (denoted with a '0') or absence (denoted with a '1') of each of the selection of diseases.

There are already a few different approaches in the problem space of thoracic disease classification from CXRs, as discussed before. We started inspired by several different models[35] [13] [15] [21], and have experimented with several different existing approaches and tweaked certain aspects to yield better results. Doing so has helped us significantly in developing a better understanding of the overall model. Experimenting with different techniques we have finally ended up with an innovative model of our own. Having tested out different existing techniques, we detected the incompetencies in the data and tried to cater to them. Our approach is designed to get the most out of the available data and manipulate it in such a way that produces better results.

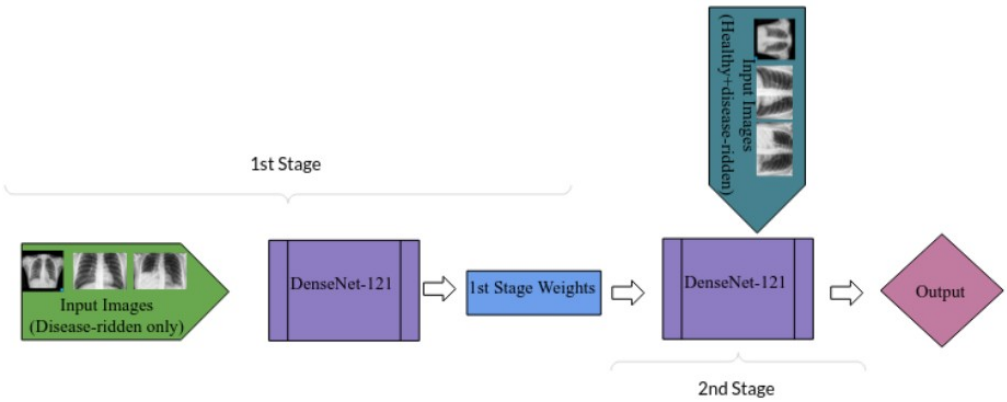


Figure 4.2: The architecture of the proposed model.

We have designed a model which works in stages and deals with some of the data at a time as portrayed in Figure 4.2. This significantly speeds up the training process. As mentioned earlier, more than half the data does not even contain disease features. As the priority of the model is to detect, classify and label the presence of a disease, the model must be very familiar with CXRs of patients diagnosed with a disease, rather than the one of a healthy patient.

Stage 1

The first stage of our model is trained solely on the CXRs of patients with a disease label, i.e. not the ones labeled "No Findings". This helps the model better recognize and learn the features of each disease, and also significantly decreases training time.

The model is designed to do this in several sub-stages, with increasing numbers of epochs each time, to further quicken the process. The entire network uses the DenseNet-121 model[16] as the base model.

We use transfer learning to start training. Instead of starting from scratch, we initialize the first stage with the weights from the ImageNet[19] model. This helps the network "understand" how to classify and label a large range of objects, and thus it can extract features from new data more efficiently. This can be analogous to how it will be easier for a college student to understand deep learning than it would be for a kindergarten student who has just started learning.

The first phase of stage 1 is initializing the model with the weights from ImageNet [19] and training the network on the non-healthy images from the ChestX-ray14 dataset[39] [28] with Adam optimizer. This optimization algorithm helps deal with the noise present in the Chest X-rays. We start with a learning rate of 0.001, and gradually increase the number of epochs and reduce the learning rate by a factor of 10 each time if there is a plateau. At the beginning of the rest of the phases, we initialize the model with the weights from the previous phase (instead of the ImageNet weights we used from transfer learning) and continue training. This concludes the first stage of the model.

Stage 2

The second stage starts from where the first one ends, using the weights the model has learned from the preceding phases. This time, however, the model uses all 112,120 images of the dataset, including the ones on which it has been trained in the first stage (the non-healthy patients' CXRs). This teaches the model how to identify the absence of these diseases as well. We repeat the phases of stage 1 in stage 2 and this time we used the entire dataset. Since there is a significant increase in the amount of data as compared to the first stage, this stage takes longer. We fine-tuned the optimizer's learning rate for different epochs. Once all phases have been completed, the model has been completely trained.

In the final layer of our DenseNet model, we use the sigmoid function, and the final output is a binary column vector, representing the predicted probability of the presence or absence of each of the 14 diseases.

4.2 Developing a Web Application

In this study, a web application, named Thoracic Complication Detector [1], was developed to classify between healthy and disease-ridden images. The user interface of the web application is shown in Figure 4.3. In the web application, a CXR image can be uploaded and based on the uploaded image, the web application provides a verdict for possible complications.

[1] https://thoracic-detection.herokuapp.com/

The proposed model was used at the back end of the web application. Whenever an image is uploaded in the web application, the trained CNN model is loaded and it predicts the class label of the uploaded image. The web application was developed by using Flask [2],which is a popular full stack web development framework written in python. At first, the web application was built on a local machine. Then, it was hosted on Heroku [3], which is a platform for running, building and operating applications in cloud.

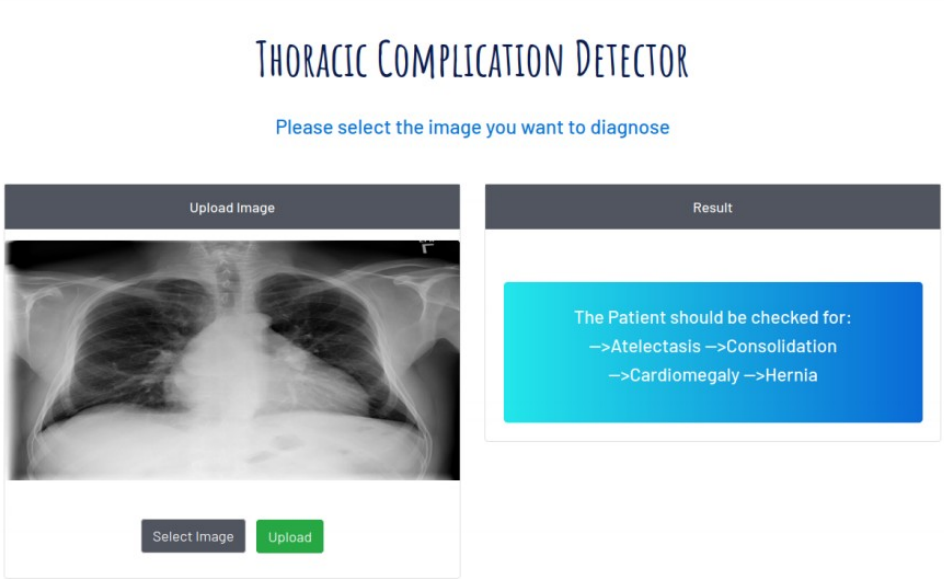


Figure 4.3: User Interface of the proposed thoracic complication detector system.

[2]https://flask.palletsprojects.com/en/1.1.x/
[3]https://www.heroku.com/

Chapter 5

Result Analysis and Discussion

5.1 Result Evaluation

While our multi-stage optimization approach to detecting thoracic diseases from CXRs starts with a low AUC score of about 0.607 during the very first phases of stage 1, after stage 2 is completed, the score rises to about 0.829. Our model performs better for more than half the diseases out of the 14 diseases and gives satisfactory results for the other 6 as well.As the ROC curve in Figure 5.1 shows, the model performs well for all 14 diseases, while performing the best for Hernia, and the worst for Infiltration.

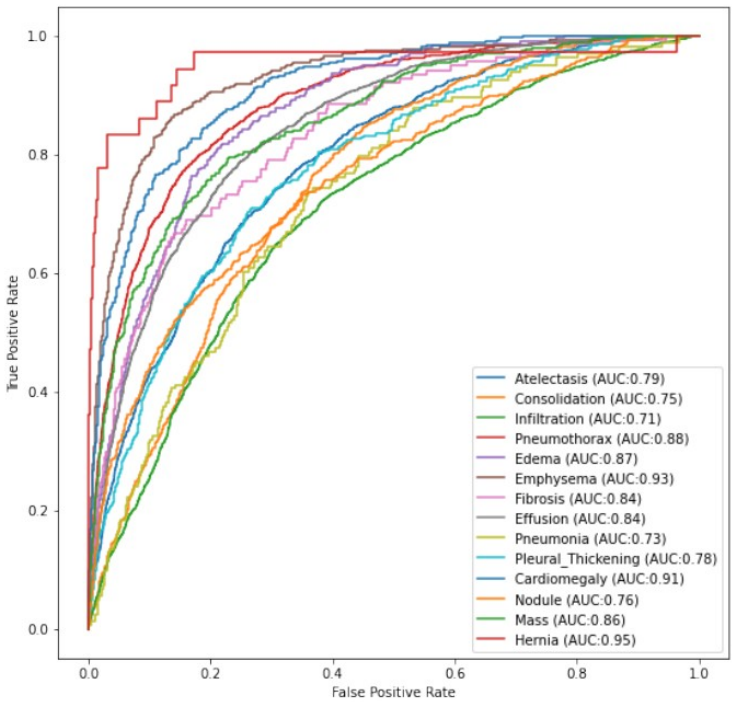


Figure 5.1: ROC curve for all the 14 diseases showing the True Positive vs False Positive rate which is indicating the model's discriminative ability.

Our model performs significantly better than previous models developed by Wang et al.[15] and Yao et al.[21] across almost all 14 diseases. We get better results for Atelectasis, Infiltration, Pneumothorax, Edema, Effusion, Cardiomegaly, Mass, Hernia

than the other 3 models. Our results are more comparable with the A^3 model[35], the state-of-the-art benchmark of thoracic disease detection models, with similar results and even beating their score in 8 diseases out of the 14 as displayed in Table 5.1. The A^3 model gives a score of 0.938 in predicting Hernia, while the model developed by Yao et al. [21] yields an even lower value of 0.914. Our model outperforms them with an AUC score of 0.950 and gives better prediction results.

Pathology	Wang et at. (2017)	Yao et al. (2017)	A^3Net (2021)	Proposed Multi-Stage Optimization
Atelectasis	0.716	0.772	0.779	0.78885
Consolidation	0.708	0.788	0.759	0.75264
Infiltration	0.609	0.695	0.710	0.71155
Pneumothorox	0.806	0.841	0.878	0.88464
Edema	0.835	0.882	0.855	0.86904
Emphysema	0.815	0.829	0.933	0.92624
Fibrosis	0.769	0.767	0.838	0.83530
Effusion	0.784	0.859	0.836	0.84499
Pneumonia	0.633	0.713	0.737	0.73181
Pleural Thickening	0.708	0.765	0.791	0.77917
Cardiomegaly	0.807	0.904	0.895	0.91113
Nodule	0.671	0.717	0.777	0.75844
Mass	0.706	0.792	0.834	0.85568
Hernia	0.767	0.914	0.938	0.95034

Table 5.1: Comparison of AUC scores of our model, versus the benchmark, popular models in thoracic disease detection.

We get significantly better prediction results for Cardiomegaly as opposed to all the other three models. Our model gives an AUC score of 0.911 while the others give 0.895, 0.904, 0.807 for Cardiomegaly. From Table 5.2 we can see the model's performance metrics for precision, recall, and F1 score of all the 14 diseases. 10-fold cross-validation is used to evaluate the model. We have taken the mean average of all the folds for a better insight into the performance. The model performs significantly well on all 14 diseases, and exceptionally well for Hernia with a recall score of 1.00 and 0.99 Precision and F1-score.

Disease	Precision	Recall	F1-score
Atelectasis	0.76	0.87	0.81
Consolidation	0.87	0.93	0.90
Infiltration	0.58	0.76	0.65
Pneumothorax	0.80	0.89	0.84
Edema	0.93	0.96	0.95
Emphysema	0.91	0.96	0.93
Fibrosis	0.97	0.98	0.97
Effusion	0.66	0.82	0.73
Pneumonia	0.96	0.98	0.97
Pleural_Thickening	0.92	0.96	0.94
Cardiomegaly	0.92	0.96	0.94
Nodule	0.88	0.94	0.91
Mass	0.87	0.93	0.90
Hernia	0.99	1.00	0.99

Table 5.2: K-Fold cross validation performance metrics of the proposed model

5.2 Scope and Limitations

This study aimed to develop a deep learning model that can efficiently diagnose up to 14 different thoracic diseases, from a CXR alone. Thoracic diseases are one of the leading causes of death in children and adults alike[40]. If implemented and further developed this model like the many others in this field is meant to aid doctors and help save lives. Having experimented with the existing approaches to solving this problem, we ended up with a multi-stage optimization model that works in 2 different stages, each fine-tuned in a way to cater to the discrepancies of the data. We managed to outperform the state-of-the-art models in this field in detecting almost all of the 14 diseases, and also achieved a high AUROC mean score of 82.9%.

However, as with any system, our model has some limitations that need to be dealt with if we were to make use of this system in the medical diagnostic field. One significant limitation of our approach is the lack of incorporation of patient history or any sort of holistic diagnosis approach to the problem. As the dataset itself does not contain supplementary information on the patient and any track of the presence of other problems in their body, there might be cases of misdiagnosis. Another limitation is the incoherent data itself, which does not provide the system to learn with consistent amounts of data across the spectrum of diseases. Some diseases have a lot more images available, and some have very few. Therefore, the model cannot properly train to the same accuracy levels for detecting all 14 diseases.

5.3 Future Work

The model as it is now, is not suitable to be used in the medical diagnosis field. Since one wrong move can affect someone's life, we need much better performance and accuracy in our predictions if we are to aid the diagnosis process. Our model

has a lot of potential and can be further developed to incorporate patient histories and underlying diseases if given the chance. However, since the original diagnosis reports accompanying the CXRs that we have used have not been made publicly available by NIH, we could not utilize them. If made available, these reports can help improve the accuracy of these prediction models greatly.

Any future work with our existing model might also compensate for the inconsistent data across the 14 thoracic diseases. If a more even distribution is possible, i.e. more data of diseases like Hernia, Edema, Pneumonia, etc are fed into the system, then the system might recognize these diseases even more accurately. However, even with better-equipped systems, medical diagnosis models built with neural networks are not meant to replace doctors, but simply aid them and speed up the diagnosis process.

Doctors with their expertise, have years of experience and have a much better understanding of these diseases and their diagnosis processes. Including a doctor or incorporating their opinion into the development phases might also greatly increase the scope of this model.

As can be seen, there is much stride left to be made before we can envision the implementation of these models in the medical field. But the high AUC scores give us hope that progress is possible. It won't be long before we will see neural network models being used alongside X-ray machines and CT scan machines.

Chapter 6

Conclusion

Pneumonia alone affects about 150 million people worldwide and can cause up to 4 million deaths in a year. These thoracic diseases, if diagnosed early and treated properly can save millions of lives. Though our model is yet to be suitable for practical use in the medical field, it is an example of how much potential neural networks and deep learning have in the advancement of medical science. Our model works in 2 stages, dividing the data into healthy and unhealthy patients and using a base model of DenseNet-121. We managed to reach an AUROC mean score of 82.9% and have yielded results better than the state-of-the-art models in this field, in 8 of the diseases out of the 14 labels. If neural network models like ours can improve the diagnosis process and make it more efficient, we can help greatly reduce the mortality rate related to thoracic diseases. We look forward to a future where this model, like many others, will contribute to a digital world in which machines can help save lives.

Bibliography

[1] A. K. Dixon, "Evidence-based diagnostic radiology," *The Lancet*, vol. 350, no. 9076, pp. 509–512, 1997.

[2] K. Lee, "Principles of cad/cam/cae systems," 1999.

[3] L. R. Medsker and L. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, 2001.

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

[5] M. Oken, W. Hocking, P. Kvale, G. Andriole, S. Buys, T. Church, E. Crawford, M. Fouad, C. Isaacs, D. Reding, J. Weissfeld, L. Yokochi, B. O'Brien, L. Ragard, J. Rathmell, T. Riley, P. Wright, N. Caparaso, P. Hu, G. Izmirlian, P. Pinsky, P. Prorok, B. Kramer, A. Miller, J. Gohagan, and C. Berg, "Screening by chest radiograph and lung cancer mortality: The prostate, lung, colorectal, and ovarian (plco) randomized trial," English, *Journal of the American Medical Association*, vol. 306, no. 17, pp. 1865–1873, Nov. 2011, ISSN: 0098-7484. DOI: 10.1001/jama.2011.1591.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[7] N. Murray and F. Perronnin, "Generalized max pooling," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2473–2480.

[8] D. Demner-Fushman, M. D. Kohli, M. B. Rosenman, S. E. Shooshan, L. Rodriguez, S. Antani, G. R. Thoma, and C. J. McDonald, "Preparing a collection of radiology examinations for distribution and retrieval," *Journal of the American Medical Informatics Association*, vol. 23, no. 2, pp. 304–310, 2016.

[9] K. Greff, R. K. Srivastava, J. Koutnık, B. R. Steunebrink, and J. Schmidhuber, "Lstm: A search space odyssey," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2222–2232, 2016.

[10] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," *arXiv preprint arXiv:1611.01144*, 2016.

[11] M. A. Makary and M. Daniel, "Medical error—the third leading cause of death in the us," *BMJ*, vol. 353, 2016. DOI: 10.1136/bmj.i2139.

[12] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, Ieee, 2017, pp. 1–6.

[13] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. Lungren, and A. Ng, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," Nov. 2017.

[14] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *arXiv preprint arXiv:1706.03762*, 2017.

[15] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017. DOI: 10.1109/cvpr.2017.369. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2017.369.

[16] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, *Densely connected convolutional networks*, 2018. arXiv: 1608.06993 [cs.CV].

[17] Q. Ke, J. Liu, M. Bennamoun, S. An, F. Sohel, and F. Boussaid, "Chapter 5 - computer vision for human–machine interaction," in *Computer Vision for Assistive Healthcare*, ser. Computer Vision and Pattern Recognition, M. Leo and G. M. Farinella, Eds., Academic Press, 2018, pp. 127–145, ISBN: 978-0-12-813445-0. DOI: https://doi.org/10.1016/B978-0-12-813445-0.00005-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128134450000058.

[18] D. Kermany, K. Zhang, and M. Goldbaum, "Large dataset of labeled optical coherence tomography (oct) and chest x-ray images," *Mendeley Data, v3 http://dx. doi. org/10.17632/rscbjbr9sj*, vol. 3, 2018.

[19] X. Wang, Y. Peng, L. Lu, Z. lu, and R. Summers, "Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays," Jan. 2018. DOI: 10.1109/CVPR.2018.00943.

[20] C. Yan, J. Yao, R. Li, Z. Xu, and J. Huang, "Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays," ser. BCB '18, Washington, DC, USA: Association for Computing Machinery, 2018, pp. 103–110, ISBN: 9781450357944. DOI: 10.1145/3233547.3233573. [Online]. Available: https://doi.org/10.1145/3233547.3233573.

[21] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, and K. Lyman, *Learning to diagnose from scratch by exploiting dependencies among labels*, 2018. arXiv: 1710.10501 [cs.CV].

[22] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, IEEE, 2018, pp. 1–2.

[23] I. Allaouzi and M. B. Ahmed, "A novel approach for multi-label chest x-ray classification of common thorax diseases," *IEEE Access*, vol. 7, pp. 64 279–64 288, 2019.

[24] K. Eckle and J. Schmidt-Hieber, "A comparison of deep networks with relu activation function and linear spline-type methods," *Neural Networks*, vol. 110, pp. 232–242, 2019.

[25] S. Gündel, S. Grbic, B. Georgescu, S. Liu, A. Maier, and D. Comaniciu, "Learning to recognize abnormalities in chest x-rays with location-aware dense networks," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, R. Vera-Rodriguez, J. Fierrez, and A. Morales, Eds., Cham: Springer International Publishing, 2019, pp. 757–765, ISBN: 978-3-030-13469-3.

[26] A. E. W. Johnson, T. J. Pollard, N. R. Greenbaum, M. P. Lungren, C.-y. Deng, Y. Peng, Z. Lu, R. G. Mark, S. J. Berkowitz, and S. Horng, *Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs*, 2019. arXiv: 1901.07042 [cs.CV].

[27] M. M. A. Monshi, J. Poon, and V. Chung, "Convolutional neural network to detect thorax diseases from multi-view chest x-rays," in *Neural Information Processing*, T. Gedeon, K. W. Wong, and M. Lee, Eds., Cham: Springer International Publishing, 2019, pp. 148–158, ISBN: 978-3-030-36808-1.

[28] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray: Hospital-scale chest x-ray database and benchmarks on weakly supervised classification and localization of common thorax diseases," in *Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics*, L. Lu, X. Wang, G. Carneiro, and L. Yang, Eds. Cham: Springer International Publishing, 2019, pp. 369–392, ISBN: 978-3-030-13969-8. DOI: 10.1007/978-3-030-13969-8_18. [Online]. Available: https://doi.org/10.1007/978-3-030-13969-8_18.

[29] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.

[30] T. Mahmud, M. A. Rahman, and S. A. Fattah, "Covxnet: A multi-dilation convolutional neural network for automatic covid-19 and other pneumonia detection from chest x-ray images with transferable multi-receptive feature optimization," *Computers in Biology and Medicine*, vol. 122, p. 103 869, 2020, ISSN: 0010-4825. DOI: https://doi.org/10.1016/j.compbiomed.2020.103869. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482520302250.

[31] S. Purkayastha, S. B. Buddi, S. Nuthakki, B. Yadav, and J. W. Gichoya, "Evaluating the implementation of deep learning in librehealth radiology on chest x-rays," in *Advances in Computer Vision*, K. Arai and S. Kapoor, Eds., Cham: Springer International Publishing, 2020, pp. 648–657, ISBN: 978-3-030-17795-9.

[32] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, and R. Agha, "World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19)," *International Journal of Surgery*, vol. 76, pp. 71–76, 2020, ISSN: 1743-9191. DOI: https://doi.org/10.1016/j.ijsu.2020.02.034. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1743919120301977.

[33] H. Wang, H. Jia, L. Lu, and Y. Xia, "Thorax-net: An attention regularized deep neural network for classification of thoracic diseases on chest radiography," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 475–485, 2020. DOI: 10.1109/JBHI.2019.2928369.

[34] W. Zhao, Z. Zhong, X. Xie, Q. Yu, and J. Liu, "Relation between chest ct findings and clinical conditions of coronavirus disease (covid-19) pneumonia: A multicenter study," *American Journal of Roentgenology*, vol. 214, no. 5, pp. 1072–1077, 2020.

[35] J. Wang, H. Zhu, S. Wang, and Y.-D. Zhang, "A review of deep learning on medical image analysis," *Mobile Networks and Applications*, vol. 26, Feb. 2021. DOI: 10.1007/s11036-020-01672-7.

[36] W. Brant and C. Helms, "Fundamentals of diagnostic radiology, 4th ed.philadelphia, pa, 2012, pp. 2."

[37] "Forum of international respiratory societies. the global impact of respiratory disease – second edition. sheffi eld, european respiratory society, 2017.," (Accessed on 30/05/2021).

[38] K. H.J. Wright and J.Inch, "The general practitioner's use of diagnostic radiology, journal of the royal society of medicine, vol. 72, pp. 88-94, feb. 1979."

[39] "Nih chest x-rays, nih clinical center, bethesda, maryland, 2018," (Accessed on 30/05/2021).

[40] "World health organization, global surveillance, prevention and control of chronic respiratory diseases: A comprehensive approach, geneva, 2013," (Accessed on 30/05/2021).