中图分类号：　　　　　　　　　　　　　　　　　　学校代码：　10055

UDC：　　　　　　　　　　　　　　　　　　　　　　　密级：　公开

南開大學

Nankai University

硕 士 学 位 论 文

Master’s Dissertation

论文题目: 使用胸部X射线图像的基于CNN的肺炎检测和预测

Title: **CNN based pneumonia detection and prediction using chest X-ray images**

论文作者 Wolde Kibruyisfa Kasu　Author Wolde Kibruyisfa Kasu

指导教师 BO ZHANG　Supervisor BO ZHANG

申请学位 软件工程硕士　Degree Master of Software Engineering

培养单位 软件学院　Institute College of Software

学科专业 软件工程　Discipline Software Engineering

研究方向 机器学习　Research Area machine learning

答辩委员会主席 张玉志　Chairman of Committee Zhang Yuzhi

评 阅 人 高铁杠 王超　Reviewers Gao Tiegang Wang Chao

南开大学研究生院

Graduate School of Nankai University

2020 年 5 月

The Authorization of Dissertation in Nankai University

I fully understand the management regulations of Nankai University on the retention and utilization of graduate student's dissertation, and agree to submit my electronic form and corresponding paper version of my dissertation to Nankai University.

I understand that Nankai University has the right to utilize the dissertation within “The Law of Copyright” of the People's Republic of China, and I agree to authorize the university in the following aspects. That is:

1. The university will adopt the dissertation into "The Full-text Database of Nankai University’s Dissertations "., and can provide the public dissertations to teachers and students as the reading in the library and free information services including catalog search, abstracts and the first 16-pages browsing on the campus network;

2. The university can adopt photocopying, micro-printing or other means to save the copy of dissertation. According to the provisions of the Ministry of Education, the university should submit public dissertations to the institution designated by the Ministry of Education;

3. This authorization is applicable to confidential dissertations after decryption.

I promise: the dissertation is completed by myself during studying in Nankai University and has been passed the defend; the electronic version of the dissertation is consistent with the contents of this paper; if these contents are different, I undertake my responsibility.

I have signed one copy of this authorization (that is one page of the dissertation), which shall be submitted to the library for retention

Author’s signature: ______________________ Date: 2020.06.18

The Author’s Information

<table>
<tr><td>Title</td><td colspan="5">CNN based pneumonia detection and prediction using chest X-ray images</td></tr>
<tr><td>Name</td><td>Wolde Kibruyisfa Kasu</td><td>Student ID</td><td>2120186041</td><td>The Date of Oral Defense</td><td>2020.06.01</td></tr>
<tr><td>The Type of Dissertation:</td><td colspan="5">Master</td></tr>
<tr><td>Institution/Department</td><td colspan="3">College of Software</td><td>Major:</td><td>Software Engineering</td></tr>
<tr><td>Email Address:</td><td colspan="5">kibruyisfakasu@gmail.com</td></tr>
</table>

Note: The authorization applies to all students’ dissertations. The author fills in the duplicate authorization and hands in the library. The non-public dissertation should be with “The Approval Form of non-public dissertation in Nankai University

The Original Declaration of Dissertation in Nankai University

I solemnly declare: The presented dissertation is my own original work under the guidance of my supervisor. All references have been stated and the dissertation doesn’t contain any others works created, published or not published. Other individuals and the collectives that contribute to the dissertation have been stated in the dissertation. I will be responsible for all the legal liability of the statement.

Author’s signature: _________ Date: 2020.06.18

摘要

肺炎是一个全球性的健康问题，每年导致数百万人死亡。肺炎的风险对许多人来说是巨大的，尤其是在发展中国家，数十亿人面临生命力下降，并依赖于污染形式的能源。在医学研究中，细菌，病毒或真菌可能引起肺部感染或肺炎。肺炎是一种严重的呼吸道感染，会影响肺部。肺由称为肺泡的小囊组成，当一个健康的人呼吸时，这些囊会充满空气。胸部X射线用于分析和诊断肺炎，需要专业的放射治疗师进行评估。因此，建立一个用于识别和检测肺炎的自动框架或系统将对治疗感染有好处，尤其是在偏远和农村地区。由于深度学习算法在分析医学图像方面取得了成功，因此卷积神经网络（CNN）在疾病分类和预测方面引起了广泛关注。

使用不同的医学影像技术检测肺炎变得非常具有挑战性，因为在检测过程中任何错误的检测都可能导致严重的医疗后果。准确的识别对于任何干预都很重要。因此，在自动检测过程中利用技术非常重要，也非常重要。

这项研究建议通过使用一系列胸部X射线图像样本训练卷积神经网络（CNN）模型来分类和检测肺炎的存在。基于胸部X射线图像的放射线图像，我们将其分类以确定一个人是否感染了肺炎。该模型可以帮助缓解在处理医学图像时经常遇到的可靠性和可解释性挑战。众所周知，与其他深度学习分类项目不同，获取医学图像数据集很困难，对于分类分配来说，找到大量肺炎数据集很困难，因此我们实施了一些数据扩充和预处理算法来进行验证和CNN模型的分类准确性，我们得到了非常出色的验证准确性。不幸的是，由于缺乏大量的数据集，并且主要是由于缺乏硬件基础架构，因此从头开始开发或构建复杂的深度指导模型并进行从头开始培训是不可行的。因此，本文提出了转移学习的思想，这意味着它正在利用来自相关预测的知识转移来改进已经学习的新预测任务中的学习。为了更有效地诊断和检测X射线图像中肺炎的存在，该技术将基于深度学习的使用来改进当前的计算机视觉方法。

关键字：机器学习，深度学习，卷积神经网络，计算机辅助诊断，儿科，肺炎，可视化，解释，胸部X射线。

Abstract

Pneumonia is a global health problem it is causing millions of deaths each year. The risk of pneumonia is enormous for many, especially in developing countries where billions confront vitality destitution and depend on polluting forms of energy. In the medical study of lungs' infection or pneumonia may be caused by bacteria, viruses or fungi. Pneumonia is a form of a severe respiratory infection that affects the lungs. The lungs are made up of small sacs like structure called alveoli, which fill with air when a healthy person breathes. Chest X-Rays are used to analyze and diagnose pneumonia and it requires an expert radiotherapist for evaluation. Hence, creating an automatic framework or system for identifying and detecting pneumonia would be beneficial for treating the infection without any delay especially in remote and rural areas. Due to the success of deep learning algorithms in analyzing medical images, Convolutional Neural networks (CNNs) have picked up much attention for disease classification and prediction.

Detecting pneumonia using different medical images techniques becomes very challenging because in the process of detecting any wrong detection could lead to serious consequences in medical treatment. Accurate identification is important to any kind of intervention. Therefore, leveraging technology in the process of automatic detection is very important and very essential.

This study proposes to classify and detect the presence of pneumonia by training a convolutional neural network (CNN) model using a collection of chest X-ray image samples. Based on the radiography image from the Chest X-ray image we classify it to determine if a person is infected with pneumonia or not. This model could help relieve the reliability and interpretability challenges often confronted when dealing with medical imagery. As we all know that getting medical images dataset is difficult, unlike other deep learning classification projects, finding a pneumonia huge amount of several data set is difficult for the classification assignment, therefore, we have implemented a few data augmentation and preprocessing algorithms to progress the validation and the classification accuracy of the CNN model and we get very remarkable validation accuracy. Unfortunately, developing or constructing a complex deep leering model and training from scratch is mostly infeasible due to the lack of a huge number of the dataset and mainly due to the lack of hardware infrastructure. Therefore, this paper proposes the idea of transfer learning which means that this is using a transfer of knowledge from a related prediction the improvement of learning in a new prediction task that has already been learned. For more effective diagnose and detect the presence of pneumonia in x-ray images this technique will improve the current computer vision methods based on the use of deep learning.

Keywords: Machine learning, deep learning, convolutional neural network, computer-aided diagnosis, pediatric, pneumonia, visualization, explanation, chest X-rays.

Contents

摘要 ... iv

Abstract ... v

List of figures ... ix

List of Abbreviations ... x

List of table ... xi

CHAPTER ONE ... 1

1. Introduction ... 1

1.1. Statement of the problem ... 3

1.2. The objective of the research ... 5

1.2.1. General Objective ... 5

1.2.2. Specific objective ... 5

1.3. Research Questions ... 6

1.4. The Significance of the Study ... 6

1.5. Structure of the Thesis ... 6

CHAPTER TWO ... 8

2. Background and Literature Review ... 8

2.1. Background of pneumonia ... 8

2.1.1. What are the symptoms of pneumonia? ... 10

2.1.2. Lung Changes Associated with pneumonia infection ... 11

2.2. Deep Learning in Medical Imaging ... 17

2.2.1. Challenges ... 19

2.4. Introduction to Deep Learning ... 21

2.4.1. Deep Learning inspired by the human brain ... 22

2.4.2. Types of Deep leaning ... 24

2.4.3. How are supervised Deep Learning algorithms trained? ... 24

2.4.4. Improving Deep Neural Network: Cases of underfitting or overfitting ... 25

2.4.5. Deep Neural Network: Optimization ... 26

2.4.6. Deep Neural Network: Regularization ... 28

2.4.7. Convolutional neural network ... 28

2.5. Transfer Learning ... 30

2.5.1. Recently Proposed Deep Learning Architectures 31
2.6. Related Work 33
Chapter Three 35
3. Data and Methodology 35
3.1. Datasets 35
3.2. Image Preprocessing and Augmentation. 37
3.3. Methodologies 40
3.4. Deep Learning Architecture VGG16 40
3.4.1. Implementation of the network 41
3.4.2. Training the Network 43
3.4.3. Overfitting in VGG16 44
3.5. Inception V3 Architecture 45
3.5.1. Preprocessing Stage 46
3.6. Model and solver definition 47
Chapter Four 49
4. Evaluation and Discussion 49
4.1. Accuracy of our models 49
4.3. Result Discussion and some limitations 55
4.4. Comparison between Our Experiment and Previous Works 56
Chapter Five 58
5. Conclusion and Future Work 58
5.1. Conclusion 58
5.2. Future work 59
Reference 60
Acknowledgments 66

List of figures

Figure 1.1. Healthy and pneumonia lung anatomy 4
Figure 2.1. Sample images without pneumonia (normal lung) 12
Figure 2.2. Sample images with pneumonia (infected lung) 12
Figure 2.3. Image with alveolar pneumonia. 13
Figure 2.4. Sample images with Chlamydophila pneumonia. 13
Figure 2.5. Sample images with Cryptococcus neoformans pneumonia 14
Figure 2.6. Traditional automatic learning flow and Deep Learning. 18
Figure 2.7. Two phases of the machine learning algorithm for classifying images. 21
Figure 2.8. Biological Neuron versus Artificial Neural Network. 22
Figure 2.9. Feedforward structure with two hidden layers 23
Figure 2.10. Basic structure of a neural convolutional network 30
Figure 3.1. Illustration or graphical representation of our data 36
Figure 3.2. Detected boundaries in sample pediatric CXRs. 40
Figure 3.3. Sigmoid logistic curve 42
Figure 3.4. An illustration of the VGG16 architecture. 43
Figure 3.5. An illustration of inception V3 model 46
Figure 4.1. (a) Accuracy and (b) loss graphics of the VGG16 model the fifth experiment in 60 epochs. 50
Figure 4.2. **(c)** Accuracy and (d) loss graphics of the VGG16 model the last sixth experiment in 60 epochs. 51
Figure 4.3. (e) Accuracy and (f) loss graphics of the inceptionV3 model the last or the fifth experiment in 100 epochs. 51
Figure 4.4. (g) Accuracy and (h) loss graphics of the inceptionV3 model the last or the sixth experiment in 100 epochs. 52
Figure 4.5. The confusion matrix for testing 624 images in VGG16 our trained model 54
Figure 4.6. The confusion matrix for testing 624 images in inceptionV3 our trained model ...54

List of Abbreviations

AI **=** artificial intelligence

ML = machine learning

NLP = natura**l** language processing

CNN = Convolutional neural network

DNN = Deep neural network

WHO = world health organization

EHR **=** Electronic Health Records

JPEG = Joint Photographic Experts Group

CAD **=** Computer aided diagnosis

CAP = community-acquired pneumonia

COPD = chronic obstructive pulmonary disease

CXRs = chest x-rays

GPU = Graphics Processing Unit

SGD = stochastic gradient descent

PCA = perform Principal component analysis

RGB = red, green, and blue

List of table

Table 2.1. Numeric example of underfitting and overfitting cases..25

Table 3.1 Dataset and its characteristics 36

Table 3.2. Settings for the image augmentation 38

Table 4.1 The accuracy of our models in each experiment and the average accuracy of all six experiments...50

Table 4.2 Network's train and test calculation time 53

Table 4.3. Based on the accuracy of the model, our results are compared with previous studies. 56

CHAPTER ONE

1. Introduction

In this chapter, we will discuss about the introduction of this thesis, the statement of the problem, then we will see the objective of the thesis, the significance of the thesis, what are the questions of this study, and then describe the structure of the thesis

Artificial Intelligence (AI) is an important field of computer science and it is thriving enormous research areas and applications. AI is an attempt by human intelligence to generate and outcome intelligent machines that can process information. Which means that simply to create or cultivate human brain-like machines. [1] AI has been part of many fields like NLP (Natural Language Processing), Expert-System, robotics, Image Processing, Image detection, etc. Machine Learning (ML) acts as a core for AI and it includes different kinds of disciplines like approximation, convex analysis, complexity theory, and probability. ML technology provides computers the ability to computations without any pre-programmed. In order to improve the performance of a computer, Machine Learning plays a vital role. [2] In the field of medical, diagnosing diseases is a challenging task. Machine Learning technology and AI have played important role in the medical fields like medical image processing, image interpretation, computer-aided diagnosis, image fusion, image segmentation, image registration, image-guided therapy, analysis techniques and image retrieval of the machine learning extract information from the image and show information effectively and efficiently. The ML and AI can help and assist doctors that they can diagnose and predict more accurate and faster the risk of diseases and prevent them in time.

When we come to interpreting X-ray images by a professional, his capacity, skill, and experience are very important, but other factors such as flurry, fatigue, errors, subjectivity regarding to group consensus opinion and long etcetera also have an effect. These variables result in significant variability among radiologists in recognizing and detecting a case of pneumonia by radiography [3]. This variability suggests that the need for a paradigm shift from the manual classification of medical images by professionals to the introduction of CAD systems (computer-aided diagnosis) to support professionals in the process of interpreting and diagnosing and also helps to minimize human error and also this technique will enhance the abilities of doctors and also for many researchers to understand that how to analyze the generic variation of that will lead to disease. In image recognition and classification task classifying medical images is the difficult tasks the main aim to classify the medical image into a category to help doctors and other medical professionals in the process of disease diagnosis or for further research.

In addition, these systems will make it possible to train new technicians with less experience and skill with reading and interpretation of chest radiographic images. [4]

Overall, medical image classification can be categorized into two steps. First, we will extract the most important features of the medical images. The next step is using the features to build a model that can classify the medical image dataset. Before AI comes, doctors and medical assistances were using their professional experience to classify and to extract features from the medical images into their different categories, which is mostly a difficult, boring, and time-consuming task. This approach likely leads to instability or no repeatable outcomes. And we should consider the great importance of medical image classification. The researchers' efforts have outcome a large number of published studies in this area. However, at present, we still cannot accomplish this mission efficiently and effectively. The classification task must be accomplished excellently because the result of this work will help in the medical sectors and also it would help doctors to diagnose diseases and also would help for further study. Therefore, how to effectively solve this task is one of the great importance.

Machine learning has generated incredible interest especially on deep learning, a machine learning branch that uses a multi-layer neural network. Deep learning is satisfactory in image classification primarily in convolutional neural networks (CNN). The main aim of this study is to develop a deep neural network by using a chest x-ray image to diagnosis and classify pneumonia In order to achieve this, convolutional neural network (CNN) along with residual networks have been employed to increase efficiency and accuracy and also to provide a very advanced introduction and practical use of machine learning and artificial neural network for medical image classification and prediction. One of the best and most accurate one for image classification is CNN-convolutional neural network. CNN applies filters to detect and differentiate certain features in images. [5][6] The working mode of CNN depends on the type of the applied filter. Therefore, when applying ML solutions to image classification, we should provide many different functions as much as possible for the network. Then after the training, analyze their values. In addition to the use of image classification in major medical fields such as diagnosis and prognosis, health insurance companies can also benefit from this researches with the increased analysis speed as well as accuracy. After all, prior and more accurate diagnosis, not to say preventive measures, ought to reduce the cost of treatment. [7]

Deep Learning is the most data-hungry technology and mostly for the medical sector there have to be more data sets to work on but within this recent years it has been a very large increase in Electronic Health Records (EHR), the improvement of Deep Learning algorithms is still hampered by the availability and accessibility of a dataset with exact annotations, a fact that's emphasized within the medical sector. Concretely in the area that concerns us, that is, the diagnosis of cases of pneumonia from chest x-ray images, we find a dataset called "Chest X-ray" from Kaggle. [8]

The chest x-ray images are prepared in 3 folders that are train, test, and Val and each folder contains two subfolders pneumonia and normal. In two class categories, all images are around 5,863 x-ray images (JPEG) [9] [10]

Chest X-ray images with both positions (anterior-posterior) were selected from retrospective cohorts of pediatric patients of 1 to 5 years old from Guangzhou Women and Children's Medical Center, Guangzhou, China. For the data set images analysis, all the images are initially screened for quality control by removing all the low quality and some unreadable scans. Then the diagnoses for the images were then ranked by two medical expert physicians before being cleared for training the AI system. In order to account for any reviewing errors, the assessment set was moreover checked by a third expert. [8][9][10]

1.1. Statement of the problem

Pneumonia is a global health problem that is not that much understanding of social or cultural strata, causing millions of deaths each year. The risk of pneumonia is immense for many, especially in developing countries where billions confront vitality destitution and depend on polluting forms of energy. The World Health Organization (WHO) estimates that over 4 million premature deaths happen yearly from household air pollution-related infections including pneumonia. as we can see from WHO data that pneumonia kills 2 million chiders under the age 5 every year and it is the leading cause of childhood mortality. In 2017 pneumonia killed almost 900,000 children at the age of 5 accounting for 15% of all deaths of the children under 5 ages. [11][12] It kills more children than HIV/AIDS, measles, and malaria combined.

Pneumonia is a form of a severe respiratory infection that affects the lungs. The lungs are made up of small sacs like structure called alveoli, which fill with air when a healthy person breathes. When an individual has pneumonia, the alveoli are filled with pus and fluid, hence the patient may feel a choking sensation, cough, and fever among other symptoms. Pneumonia is one of the single largest infectious diseases that is a cause of death in children worldwide. Pneumonia affects all peoples that can affect children and families everywhere but it’s most dominant in South Asia and sub-Saharan Africa. [12] It is noticeable that nearly all cases, approximately 95%, of clinical pneumonia occur in developing countries so, therefore, accurate and timely detection is mandatory. Children can be protected from pneumonia, it can be avoided with simple intercessions, and treated with low-cost, low-tech medicine and care.

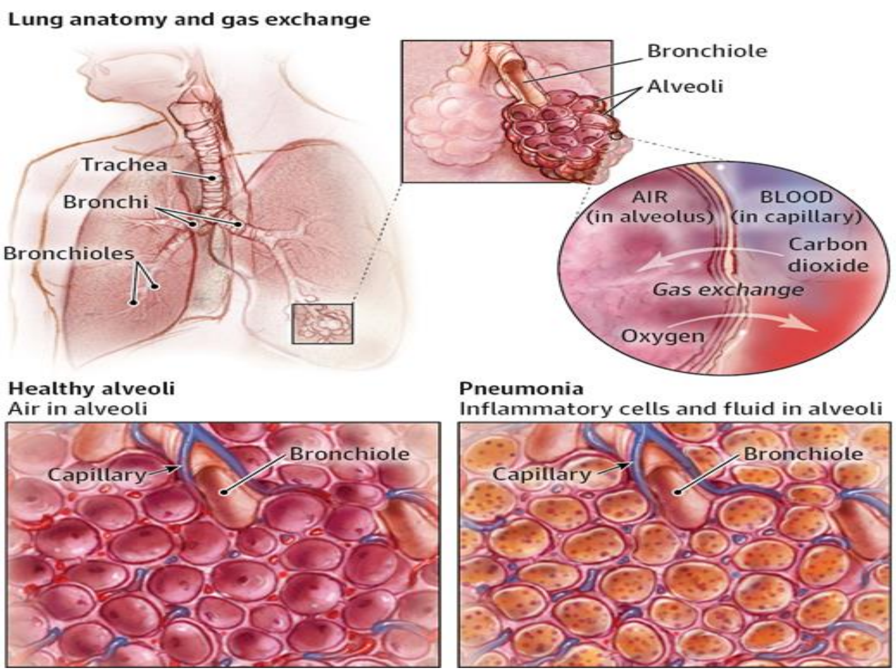


Figure 1.1. Healthy and pneumonia lung anatomy

Pneumonia is a lung infection mostly caused by bacteria, viruses, or fungi. And some Risk factors for pneumonia include age over 65 or under 2, having certain chronic medical conditions for example (including a weak immune system, underlying lung disease, smoking cigarette, alcoholism, and neurological problems), or accidental injuries that interfere with swallowing or coughing. [13]And Because of the infection and the body's immune system response, the sacks in the lungs (termed alveoli) are filled with fluids instead of air. The reason for the pneumonia chest x-ray lung opacities look diffuse is because of the infection and the fluid that accumulates spread within the normal tree of airways in the lung and this cause on the chest x-ray shadow like structure.

1.2. The objective of the research

1.2.1. General Objective

The main objective of this research is to build an implementation based on Deep Learning that allows, from chest x-rays, to detect and classify pneumonia and providing a complete diagnostic tool for cases of pneumonia. We will work in-depth on the recently created dataset "Chest X-ray" from Kaggle. Our chest x-ray images are prepared in three folders (train, test, Val) and each contains subfolders for each image category (normal and pneumonia). There are around 5,864 chest x-ray images with their class categories (normal or pneumonia). This dataset will provide us with the annotations data needed to train and evaluate our prediction and classification architecture. Throughout the development of this work, we will build a preprocessing technique, classification techniques, and detection techniques.

1.2.2. Specific objective

- Implementation and evaluation of a classification architecture based on deep neural networks.
- To clarify the nature of pneumonia and the process of disease diagnosis
- To build and train as well as test the performance of the model.
- To come up with a model that is capable to classify the accurate results of pneumonia.
- To do a deep literature review and understand what is pneumonia and how to detect pneumonia.
- Understanding the implementation of intelligent algorithms and other image processing algorithms in their usage for the detection of pneumonia.
- To design an artificial neural network pneumonia image processing.

- To compare and validate the obtained results in the study with other intelligent and other image processing methods for the detection of pneumonia.
- Evaluating and presenting the results we obtained
- Finally to forward recommendations based on the findings of the study.

1.3. Research Questions

The main question to answer is whether a solution based on deep learning can be used for the diagnosis of pneumonia disease from image data. In addition, the strategy will be evaluated to investigate how satisfactory is the proposed solution for an Implementation in essential care and hospital. The final question to answer is whether similar solutions have already been created and if there are similar works investigating their level of success and accuracy. In order to gain more deep knowledge inside the field of deep learning, recent research where deep learning is applied to different medical applications, including the treatment of pneumonia infection, will be studied.

1.4. The Significance of the Study

This study attempts to classify pneumonia using Chest x-ray image data as related to Radiology imaging attributes this can help doctors and all medical professionals use this classification method in their daily work. Although the output of this study it is very helpful for doctors to diagnose and classify pneumonia infection disease to obtain the exact results of pneumonia detection. In addition, the results of the study were used to design suitable training programs and to obtain complete classification results.

The results of this study should also serve as a reference point for researchers, medical experts and a source of methodological methods for studying image classification and detection for other similar fields of machine learning.

1.5. Structure of the Thesis

The structure of this thesis is organized as follows. Chapters 2 includes background information and Literature Review. First, we will introduce the overview of pneumonia, Chest X-ray scanning and an overview of machine learning and deep learning frameworks (such as Tensorflow), and recently proposed deep learning architectures. At the end of chapter 2, we will see some related concepts, previous related work. Chapters 3 is the main parts of this study. In Chapter 4, we describe our datasets, data preprocessing, data acquisition, and the deep learning model (VGG16) that we applied to our dataset. And we will also discuss and describe the experiments implemented for our datasets. Chapter 4 we will describe our experimental results of using the different architectural system. And finally on Chapter 5 is the conclusion of our thesis and also includes a discussion of what we will expect in the future work.

CHAPTER TWO

2. Background and Literature Review

In this chapter, we will discuss the overview of pneumonia and the diagnosis of pneumonia. Then we will describe an overview of Chest X-ray Radiology and machine learning, and a description deep learning of frameworks (such as Tensor flow), and different convolutional neural network architecture.

2.1. Background of pneumonia

Around (460– 370 BC) Pneumonia for the first time described by Hippocrates [16]. And in 1819 its clinical and pathological features were made by Laennec after the 22 century [17] in 1842 Rokitansky for the first time differentiates Lobra and bronchopneumonia. During the next 47 years at least 28 terms were used to identify pneumonia [19], and in 1929 the number had grown to 94 to the total listed in the manual of the international list of causes of death with 12 sub-terms [20]. The International Classification of Disease (ICD-10) pneumonia listed as the first and primary term in the seven codes (J12-18) classification of diseases has removed some of the historical descriptive terms but it is also a descriptive term in seven other codes as a specific infectious and non-infectious with the times of life and complications of diseases and procedures [21]. ICD-10 codes most of the time has subcategories that give pneumonia many classifications. And also we should know that some other respiratory infections include three other codes (J20–22) for acute bronchiolitis, bronchitis, and unspecified conditions.

Pneumonia is a very serious lung infection and mainly caused by bacteria, virus, or fungi and it affects one or both sides of the lungs that leads the air sacs, or alveoli, of the lungs to fill up with fluid or pus. Symptoms may include a cough with phlegm (a slimy substance), fever, chills, and trouble breathing. [14] the most common type of pneumonia is community-acquired pneumonia (CAP) this kind of pneumonia affects somebody in the community who have not recently been in the hospital or another health care facility such as a nursing home or rehab facility. The bacteria called streptococcus is the most common cause for community-acquired pneumonia but also there are many other causes. Community-acquired pneumonia is much less contagious than other infectious diseases like for example flu or a cold because most people's immune systems can kill the bacteria that cause it before they can cause an infection. Especially those who exercise more often. [15]

- **Hospital-acquired pneumonia** is a kind of pneumonia that is developed while you’re in the hospital being treated for another condition or having an operation or people in intensive care on breathing machines are at most risk.
- **Viral pneumonia** is caused by a virus and it affects the lung and the most common cause for viral pneumonia is flu, but also there are other causes for the viral pneumonia-like common cold and other viruses. These nasty germs and viruses usually stick to the upper part of our respiratory systems and by the time when they get to the lung then the trouble starts. This kind of pneumonia is contagious and can spread to others easily.
- **Aspiration pneumonia** this kind of pneumonia usually caused by food going down the wrong way, or inhaling vomit, or inhaling a foreign object or harmful substance. This one usually commonly appear on the elderly, and people who have conditions that cause swallowing difficulties or reduced level of consciousness.
- **Fungal pneumonia** is caused by fungi. Most of the time it can be caused by either endemic or opportunistic fungi or a combination of both. The mortality Case of fungal pneumonia can get high as 90% in peoples with weak immune system patients. You might also hear the term ‘double pneumonia’. This means when you get pneumonia in both lungs. It’s a term used in America.

Such as age, smoking, some medical condition, and some other factors can increase the chance of getting pneumonia and having more severe pneumonia. [14] Pneumonia affects all ages but most of the time two age groups are at a greater risk of affected by pneumonia. children under the age of five because their immune systems are still developing during the first few years of life, adults over the age of 65 because their immune systems begin to change as a normal part of aging, people with certain conditions such as diabetes, heart failure, or COPD (chronic obstructive pulmonary disease), and most of the time people who have weak immune systems because of HIV/AIDS, and peoples who pass through chemotherapy for treatment of cancer, or organ or blood and marrow stem cell transplant procedures.

2.1.1. What are the symptoms of pneumonia?

Pneumonia is caused by many different microbes like viruses, bacteria, and fungi, so if you infected with pneumonia once does not protect you from getting it again. If you are infected with pneumonia more than once you may need to have more checkups and investigations to understand it more why this is happing. It may be due to your immunity system or maybe a problem in your chest and you may be referred to a specialist. If you have pneumonia, the symptoms are similar to the flu or a chest infection. And usually, these symptoms may develop through time over a few days but can progress much faster. Coughing is the main symptom for pneumonia and also you may feel generally unwell, tired and weak and you'll probably have a chance of experiencing at least one of these symptoms too [14] [15].

- Have a high fever with hade ace
- seeing shaking chills
- coughing with mucus that may become yellow or green
- a high temperature – with sweating and shivering
- difficulty breathing
- chest pain or discomfort
- loss of appetite

If you have pneumonia, you also experience other symptoms, like nausea (feeling sick to the stomach), diarrhea and vomiting. And some people get a sharp pain around their chest while they breathe in and out. This is because of the thin lining between the lung and ribcage, which is called the pleura, is infected and inflamed. This inflammation is called pleurisy which will stop your lungs moving smoothly as you breathe. Symptoms may vary in certain peoples. Sometimes Newborns and infants may not show any signs of the infection. Or, they may have a fever, vomit, and cough, or appear restless, tired, or sick and without energy.

Older adults and peoples with serious illness or week immune systems also may have some few milder symptoms. And sometimes they may even have a lower than normal temperature. If they already have a lung disease, it may get worse.

Complications

Often, some people who have pneumonia can be successfully treated and do not have complications.

Possible complications of pneumonia may include:

Bacteremia and septic shock. Bacteremia is a serious problem in which is the bacteria spread from the initial infection site into the blood. It may lead to septic shock, a potentially fatal complication.

Lung abscesses. Lung abscesses can be treated with antibiotics. And sometimes surgery or drainage with a needle to remove the pus.

Pleural effusions, empyema, and pleurisy. There are a serious problem and pain can occur if pneumonia is not treated. Pleura is a membrane that made from two thin layers of tissues one that wraps the outside of your lungs and the other one lines the inside of the chest cavity. Pleurisy happens when the inside and the outside thin layers become irritated and inflamed and causing pain during by the time when you breathe in. The space between the two pleura is a very thin space. Pleural effusions are the cause of fluid in the pleural space. We called it empyema if the fluid becomes infected. If this happens, you may need to have surgery to remove the fluid.

Renal failure

Respiratory failure

2.1.2. Lung Changes Associated with pneumonia infection

The main role of radiography and chest x-ray imaging is to confirm the diagnosis of pneumonia. If a patient has clinical symptoms that suggest infection of pneumonia-like fever, cough or sputum, and if the imaging findings are similar to pneumonia characteristics, then the diagnosis of infectious pneumonia must be made. The examination of the images plays a vital role in classifying the infected with non-infected peoples. And based on the findings it is very useful for evaluation and treatment. Generally, it is very difficult to determine specific pathogens of infectious pneumonia based only on the chest x-ray image findings. However, the characteristics of the image findings of several pathogens have been reported and it will help to choose subsequent examinations or first antibiotics. This is especially true and use full for the exclusion of tuberculosis, which requires quite different treatment strategies from those of ordinary infectious pneumonia. The main characteristics that we can see on the images of the chest x-ray with pneumonia are like extra shadows that can cover the heart border or loss of the diaphragm with extra shadows the loss of the costphrenic angle if we cannot see the border of the heart, the diaphragm and costphrenic angle that can be the sign of pneumonia. Suggesting possible image diagnoses of specific pneumonia on chest x-ray image examinations helps determine the initial treatment. Chest x-ray Imaging examinations is very important for differentiating noninfectious diseases from infectious pneumonia. As the x-ray image findings of noninfectious diseases have extensively been investigated, they may provide enough information to suspect another kind of disease although direct comparative studies between infectious pneumonia and noninfectious diseases are limited. The examination of the images may also reveal underlying diseases that result in pneumonia or other complications. Chest radiography imaging is most of the time enough to confirm the diagnosis of pneumonia and to evaluate treatment effects, whereas computed tomography (CT) is required to suggest causative pathogens, to exclude noninfectious pneumonia and to reveal underlying diseases.

Figure 2.1. Sample images without pneumonia (normal lung)

Figure 2.2. Sample images with pneumonia (infected lung)

As we can see in figure 2.3 this is one type of pneumonia showing alveolar pneumonia in a man in his 80s. A: as a Chest radiograph shows there are an extra shadow and nonsegmental consolidation in the right middle lung field, which is determined by the minor fissure (arrow); B suggestive of upper lobe pneumonia, C: Thin-section computed tomography (B) and image (C) coronal reformatted shows a nonsegmental consolidation with air Broncho grams suggestive of alveolar pneumonia (arrows).

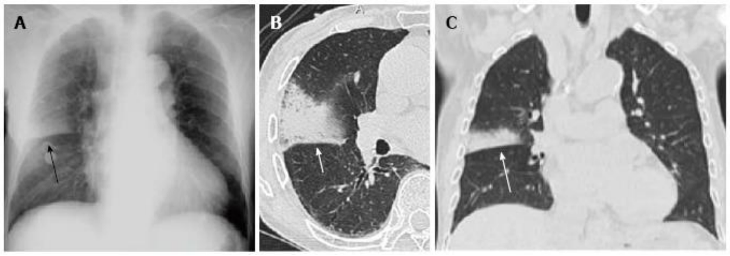


Figure 2.3. Image with alveolar pneumonia.

As we can see in figure 2.4 Chlamydophila pneumoniae showing pneumonia infectious bronchiolitis in a woman in her 60s. A: Chest radiograph shows (arrows); faint reticulonodular opacities in both lower lung fields B: Thin-section computed tomography reveals centrilobular nodules (arrows) with bronchiectasis (arrowheads) in the middle lobe and lingual.

As we can see in figure 2.5 Cryptococcus neoformans pneumonia in a woman in her 50s. A: the arrow shows on the Chest radiograph there is a mass in the right lung base B: the arrow shows Chest computed tomography with a 5 mm slice thickness shows a mass and nodule in the right lower lobe. Many nodules/masses in the same pulmonary lobe are considered characteristic findings of pneumonia.

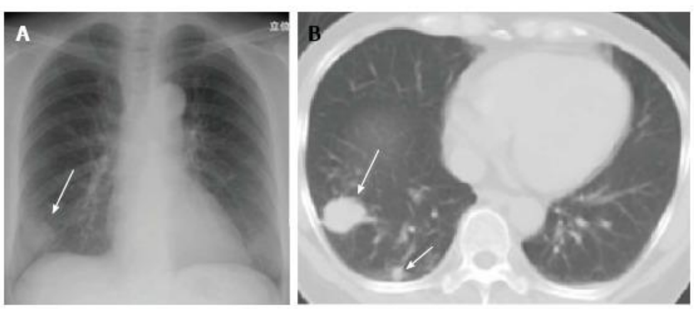


Figure 2.4. Sample images with Chlamydophila pneumonia.

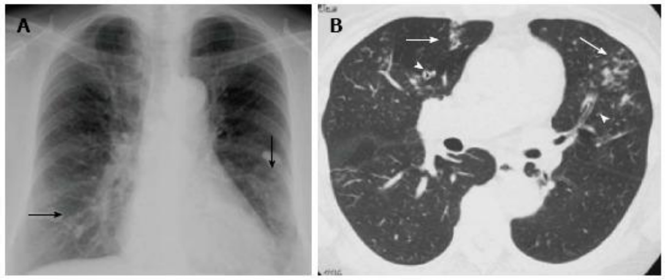


Figure 2.5. Sample images with Cryptococcus neoformans pneumonia

Radiography and chest x-ray

Chest x-ray to produce pictures it uses a very small dose of ionizing radiation to see the inside the chest. Chest x-ray usually used to evaluate the lungs, heart and chest wall and it used to diagnose the problems with breathing, fever, persistent cough, chest pain or injury. And also it is manly can be used to help diagnose different kinds of lung conditions such as pneumonia, emphysema or cancer. The chest x-ray is the fastest and easy way because of this it is particularly useful in the emergency diagnosis of treatment. A chest x-ray exam requires some preparations like not special preparation. For example, you should tell your doctor or the assistance technologist if there is a possibility you are pregnant. And don't wear any jewelry at the time of examination, and be ready to remove your clothes or you may be asked to wear a gown.

What is a Chest X-ray (Chest Radiography)? A chest x-ray is the most commonly used treatment or diagnostic x-ray examination. A chest x-ray can give us the image of the lungs, heart, airways, bones and the blood vessels and also it can show us the bones of the spine and chest. An x-ray or radiograph is a noninvasive medical examination test that is very helpful for physicians to diagnose and treat different medical conditions. Imaging with x-rays or imaging with radiograph it involves exposing a part of the body to a small dose of ionizing radiation to produce different imaging of the inside of the body. For medical imaging, X-rays are the oldest and most frequently used form.

Most of the time chest x-ray imaging is performed to evaluate the lungs, heart and chest wall.

A chest x-ray is typically the first imaging test used to help diagnose symptoms such as:

- breathing difficulties
- fever
- chest pain or injury

- a bad or persistent cough

Physicians use the chest x-ray examination to help diagnose or monitor treatment for conditions such as:

- pneumonia
- emphysema
- heart failure and other heart problems
- the positioning of medical devices
- lung cancer
- fluid or air collection around the lungs
- other medical conditions

How should I prepare?

A chest x-ray requires some preparations.

You may ask to remove jewelry, eye-glasses, metal objects, removable dental appliances, and some clothing that can affect the x-ray images. And you will be asked to remove clothes and wear a gown during the examination.

During the x-ray examination if there is a possibility of pregnancy Women must always inform their doctor and x-ray technologist. Many imaging exams will not be done during the pregnancy to minimize exposing the fetus to radiation is not good for the babe. If an x-ray is necessary, there should be taken very precautions to minimize radiation exposure to the baby.

How does the procedure work?

Chest radiography or X-ray is a form of light or radiation or radio waves that can passes through most objects including our body by carefully aimed at the part of the body during the examination and it produces a small burst of radiation that can pass through any objects and our body to take the picture or photographic film.

During the examination, our body takes or absorb the x-ray in different degrees. Different parts of the body absorb the x-rays but Dense on the bone absorbs much of the radiation while soft tissue, such as muscle, fat, and organs, allows the x-ray to pass through them. That’s why on the result, bones appear white on the x-ray, air appears black and soft tissue shows up in shades of gray.

On a chest x-ray, the ribs and spine will absorb much of the radiation and appear white or light gray on the image. Lung tissue absorbs little radiation and will appear dark on the image.

Most x-ray images are digital files which is stored as electronically. These electronically stored images are easily accessible for diagnosis and disease management.

How is the radiography procedure performed?

Usually, the technologist will position the patient in two side views of the chest one from the back and another one from the other side of the body as the patient stands against the machine or image recording plate. The technologists each and individual specially trained to perform this examination will instruct the patient with hands-on-hips and chest pressed against the image plate. For the second position view, the patient's side is against the machine or the image plate with arms elevated. And for those who cannot stand they can lay down on the table for the chest radiography.

You may be asked to hold very still and also hold your breath for a few seconds while the x-ray imaging is engaged to reduce the possibility of a blurred image, then the doctor will go the other room to start the machine.

After the examination is done, you may be asked to wait for a few seconds until the doctor tell you that all the necessary images have been taken.

Most of the time the examination process takes 15 minutes.

Additional more side views may be needed within hours, days or months to evaluate any changes in the chest.

Who interprets the results of the x-ray?

A radiologist or a professional technician specially trained to interpret and supervise the radiology examinations will analyze the images that are taken by the machine and send a report with a signature to the referring physician who will discuss the results with you.

The results of a chest x-ray will be available almost immediately for review by your physician, and if there needed a Follow-up exam your doctor will tell you because most of the time follow-up exams are needed to make sure the treatment is working further evaluations are important to see if there is a potential abnormality evaluate additional views or for a special imaging technique.

What are the limitations of Chest Radiography?

The chest x-ray is a very useful and important examination, but it has some limitations. Like for example, some conditions of the chest may not be detected with conventional chest x-ray images, chest x-ray imaging sometimes cannot necessarily rule out all the problems in the chest. For example, a blood clotting in the lungs may not show up on a chest x-ray, small cancers may not show up on a chest x-ray. A condition called a pulmonary embolism, cannot be seen on chest x-rays.

To detect this kind of conditions further imaging studies should be necessary to clarify the results of a chest x-ray or to look for abnormalities not visible on the chest x-ray.

2.2. Deep Learning in Medical Imaging

Nowadays, the need for an accurate diagnosis and prediction on medical images in order to reduce human error in the medical sector has boosted the development and improvement of computer-aided detection (CADe) and diagnosis (CADx) systems that can provide huge support and assistance to medical professionals. Moving from a scenario where the medical professionals used to diagnoses based on their knowledge and experience, and now we moved to a new scenario where the professional supports the diagnosis with systems trained in a particular problem. In the development of these health care systems different papers support the distinction of two stages [22, 23]:

- **Pre Deep Learning era**. The first computer-aided detection (CAD) is created in the early 2000s. They were built using traditional methods of Machine Learning they were systems in the process of the manual way they selected and extracted the features based on the criteria that the engineer considered appropriate. The correctness of the selection and extraction has a huge role on the system performance. Early reports warned that these systems have worse results than human readings because of the more false positives and thus and it has degutted with a question the benefit of such systems [24, 25].
- **Post-Deep Learning era**. Deep Learning and traditional machine learning techniques have many differences. Without introducing manual coding rules or human domain knowledge it can automatically learn representations from data such as images, video or text data. Deep leaning would bring a revolutionary breakthrough by discovering complicated

patterns and features from the given input data. To overcome the limitation and poor results observed in previous CAD systems the selection and extraction of the feature set without human intervention provided numerous advances and the results improved with deep learning. [4]

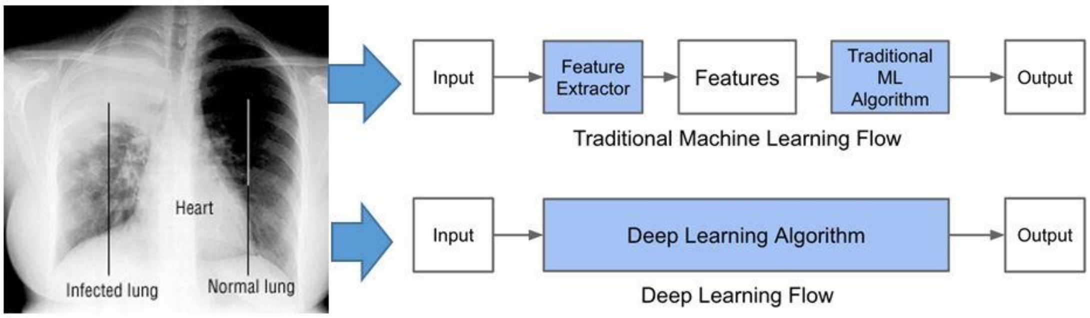


Figure 2.6. Traditional automatic learning flow and Deep Learning.

The Figure 2.6 the paradigm shift shows that the traditional machine learning usually applied according to the features that are selected and extracted manually, while the deep neural network learns the features and the most extract and most representative and complex features and patterns on its own according to the input.

The research and all the developments of Deep learning technology have advanced to next level, by extending its application in many deferent areas and in many different problems, in which it has been positioned as the most powerful and useful models as for different purposes like for example in tasks of natural language processing, voice recognition, robotic, computer vision or other recommendation systems. In the computer vision, an important breakthrough was reached when in 2012 on the ImageNet large scale visual recognition challenge a deep learning model based on convolutional networks surpassed the performance obtained with other machine learning technologies. This event confirmed the enormous potential of Deep learning technology in image processing tasks. Further developments and studies placed deep learning techniques based on convolutional networks as the effective standard for the wide variety of computer vision tasks like for example object detection, image super-resolution reconstruction, image classification, face recognition or image segmentation. The great use of deep learning has had its reflection in the CAD processing system in the medical sector providing a valuable diagnostic tool and assistance to the medical sector professional.

This momentum and success have supported in a number of factors:

- Increased availability of different digitized data.
- By increasing the powerful GPUs increasing the processing capacity
- A fast-growing development of algorithms and techniques that have significantly improved the technology that was initially proposed.
- creating and distributing of the most user-friendly software's like Keras and Pytorch
- The investment and the economic drive in the medical imaging sector based on deep learning.

2.2.1. Challenges

Several studies and researchers pass through a series of challenges for a successful implementation of Deep Learning for medical imaging tasks and act as a barrier that delays this process.

Dataset:

- The lack of a huge number of training datasets with precise annotations. This circumstance is accentuated in several pathologies less common.
- The data sets are not completely similar it is collected from different measuring devices or sensors with their corresponding calibration.
- Despite a drive to digitize medical data and implement PACS-like frameworks. There is no universal and standardized consensus of such systems and terminologies, which can lead to workflow incompatibilities.
- Dataset noisy and class imbalance. A small number of samples are available for some pathologies compared to other pathologies or cases without pathologies. This may result in an imbalanced dataset.
- The complexity involved in interpreting medical imaging requires expert and experienced staff to create the ground-truth data, with the cost of expert staff time involved.

Technological advances:

- The need for high power computing system was one challenge Years ago, this factor limited the implementation of Deep Learning systems for many years. Today, the emerging of GPU systems for increasing the power and the computational downtime techniques have partially overcome these challenges.

- Difficulty to train. As we know Medical imaging is very complex in the interpretation process and in certain cases data limitation, and to minimize the error an overtraining can leads to overfitting of the system, losing generalization to new data. In this case, we should develop the optimization and regulation techniques that have been developed and improved to minimize the effect of the overfitting.

Ethical and legal:

- Assignment of responsibilities in the event of a patient diagnostic error produced by the Deep Learning model.
- The design and use of a dataset involve legal issues. It is necessary to completely anonyms the information and in the case of images, this process can be difficult.
- Good reception and acceptance of such technology among medical professionals is necessary.
- Implementation of Deep Learning can be seen as a black box where the complex relationships and interpretations deduced from input data are opaque producing a huge barrier of understanding these complex models by clinicians.

2.3. Machine learning

Machine learning plays an important role in today's AI and technological challenges. Machine learning helps the human technological advance capability in different areas such as in the e-commerce sector, in social networking, in AI automobiles, smartphone design, cameras, in image classification, language processing, face detection, target detection, and speech recognition. Image classification has become the main case for implementing machine learning. In the machine learning algorithm for classification, there are two phases to classify image, text or data [6]. As we can see from figure 2.7 the machine learning phases for classification. The first step is the training phase by using the input images the machine learning algorithm will start training the model. In the classification phase, feature extraction is very important for training the network to extract new features from the images. There are two main feature extraction algorithms in image classification SIFT (scale-invariant feature transform) and HOG (directional gradient histogram). After we train the model then second step is the prediction or the test stage, by using the trained classifier model in the previous stage by Appling new test images previously unseen images and predict the label of the images. For the new images, we can apply the same feature extraction algorithm that is applied in the previous stage.

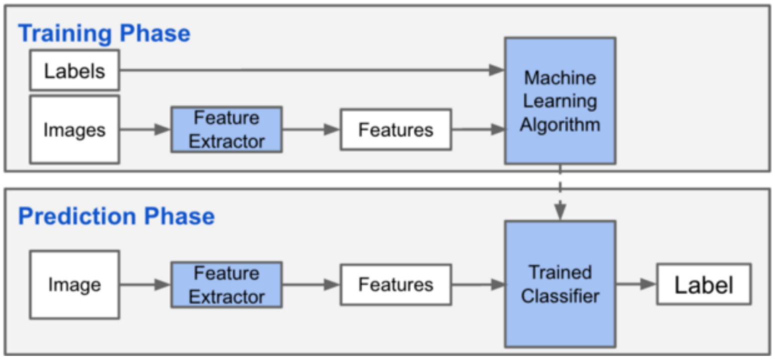


Figure 2.7. Two phases of the machine learning algorithm for classifying images.

2.4. Introduction to Deep Learning

In this section, we will briefly see the fundamentals of Deep Learning. Then we will see more the practical view introducing the most relevant aspects of the training process and optimization of neural networks. And at last, we will focus on convolutional neural networks.

Deep learning is the reason for many recent breakthroughs in artificial intelligence technologies, such as Google self-driving cars, DeepMind's AlphaGo, intelligent voice assistants and so on. Although Deep learning is a very fantastic algorithm with GPU accelerated deep learning framework so that many researchers and data scientists can significantly accelerate Deep learning training, otherwise, it may take so much time for training or days, or it may take days. When the model is ready to deploy the researchers or the data scientists can rely on the GPU accelerated cloud reasoning platform or embedded device or autopilot to provide high performance for computing the densest deep neural network.

Deep learning algorithms use a huge number of data and the lots of computing power of GPU to learn information directly from text data, or image or signal. Deep learning technology provides the flexibility to design and train custom deep neural networks and it has the interface of general programing language. For developers, Nvidia deep learning SDK provides powerful libraries and many different tools for developing deep learning framework such as a cognitive toolkit, caffe2, MXNet, Tensorflow, PyTorch, etc. Nvidia Deep learning SDK includes libraries for deep learning of primitives, reasoning, linear algebra, video analysis, sparse matrices, and multi-GPU communications. [7]

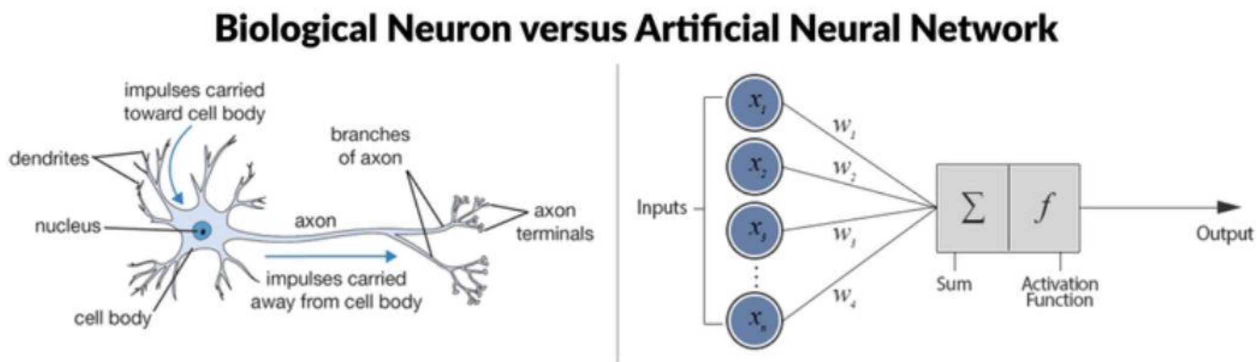


Figure 2.8. Biological Neuron versus Artificial Neural Network.

2.4.1. Deep Learning inspired by the human brain

Deep Learning is a branch of Machine Learning. And what makes the two of them similar is that their ability to automatically learn representative functions and features from the given input data, and we can call it "representation learning" or "feature learning". As we can see in figure 8 Deep Learning and human nervous system has some similarity deep learning performs out the training and learning it through multiple nodes or neurons distributed in multiple layers, simulating the functioning of the human nervous system, this similarity is shown in Figure 8. In the case of artificial neural networks (ANN), we start from a basic processing unit called node or neuron. Each neuron is itself a simple classifier model that generates an output as a function of the evidence of the previous layers. Each neuron performs three actions:

- **Linear operation.** Linear operation is on a propagation function adds the signals coming from the previous layer with their respective weights and a bias value.
- **Non-linear operation.** Non- linear operation is on the propagation function applies an activation function which determines whether or not the output of the neuron is activated. There are several activation functions and depending on the space occupied by the node

within the neural network it is convenient to use one or the other. The most common activation functions are Sigmoid, Tanh, softmax, and ReLU.

- **Output.** The result of the activation function is located at the exit of the node and will propagate to all nodes located in the next layer.

These nodes or neurons are organized in layers and through the interaction with other neurons distributed in different layers allows us to obtain more complex and hierarchical representations. Each layer essentially acts on the characteristic constructions of the layers prior to it generating a higher-level feature construction.

In the neural network, we come across "N" nodes by layer and "C" different layers. In the simplest structure, the nodes of the "C-1" layer are all connected to all the neurons of the "C" layer, this structure is called "feedforward" or "Multilayer Perceptrons" (MLP) and is the basic structure of Deep Learning. Figure 2.9 shows the feedforward structure. [4]

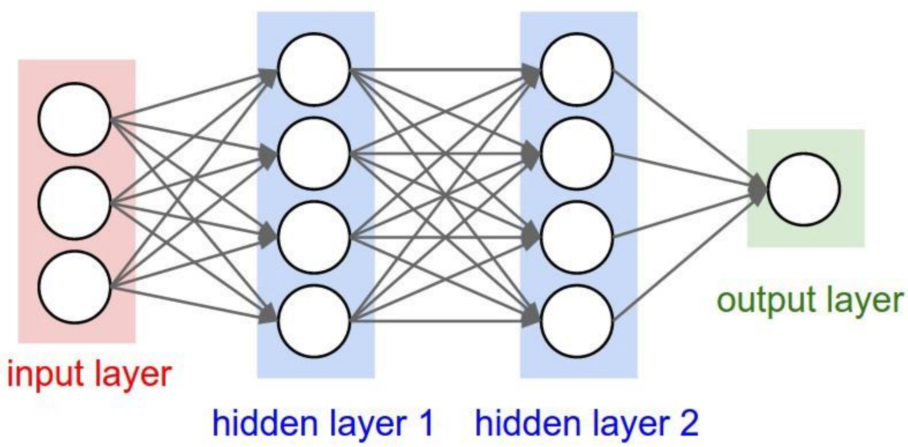


Figure 2.9. Feedforward structure with two hidden layers

There are three main parts in the neural network the Input layer, the hidden layer, and the output layer. The number of input and output layers is fixed and their shape is based on the input dataset and the formulation of the problem or target function searched for respectively. But the size of the hidden layers can be variable. A neural network that uses a single hidden layer and introduces more layers we call it “Shallow neural network” and we will have a deeper model.

2.4.2. Types of Deep leaning

There are many different types of Deep learning according to their purpose and type of data they use in the process of training. We find different algorithms, the main types are:

Supervised algorithm. This is one of the main type's deep learning and it simply means various algorithms generate a function that maps inputs to desired outputs. The model takes labeled data and trained to minimize the cost function by comparing the prediction obtained by the model with the real value. In this case, we find (FFNN) Feedforward neural network that we have mentioned previously, the Convolutional Neural Networks (CNN) used mainly in computer vision and Long Short Term Memory (LSTM) for applications based on time series data.

Semi-supervised algorithm. This algorithm trains itself by generating its own annotations [21]. Only a small part of the samples have the annotations necessary to train the algorithm. This algorithm builds a self-learning system and the algorithm itself generates its own annotations [21]. Within this type we find Deep reinforcement learning (DRL) and Generative Adversarial Network s (GAN).

Unsupervised algorithm. When you need to supervise the model it will be unsupervised learning is a machine learning technique. Or Instead, when the model to work on its own to discover information. Most of the time it will deal with the unlabeled data.

Compared to supervised learning Unsupervised learning algorithms allow you to perform more complex processing tasks. Although unsupervised learning compared to other natural learning methods it can be more unpredictable.

2.4.3. How are supervised Deep Learning algorithms trained?

In deep learning, the model learns by itself it is not necessary to define a rule of prediction and selection of the characters. As we have already indicated before. To optimize the weight of the

nodes the optimization algorithm used is called gradient descent until finding a local minimum of a function, i.e., minimize the cost function? The 3 steps of descent gradient method:

- Forward propagation. As we understand from the name, the input data is fed in the forward direction to the network. And each of the hidden layer accepts the input data, processes it with the activation function and give it to the successive layer. Based on their current weights of each node a prediction will be made.

- Evaluation. The obtained prediction is compared with the objective value, obtaining an estimated value from the error of the configuration of the current weight in the nodes.
- Backward propagation. This is an error function given artifice this error is propagated in the opposite direction from the output to the input by applying slight changes in the weights proportional to the error and in the direction that minimizes the error. With respect to the neural network's weights an artificial neural network and an error function, the method calculates the gradient of the error function. [4]

2.4.4. Improving Deep Neural Network: Cases of underfitting or overfitting

The main objective and goal of training a deep learning algorithm are that the model acquires the ability to define or generalize what you have leaned, that is from the data used in the training phase and be able to predict accurately. The loss of generalizability of the model can occur in two directions [26]:

• Underfitting. Happens when not been properly trained or the structure of the model is too simple to create an accurate representation of the input data. Showing a high bias, i.e. the predicted value and the actual value is high and when there is a difference between values.

• Overfitting. When the model has been overtrained and the nodes have ended up "memorizing" the input data. Showing a high variance, i.e. a significant difference between the error of the train set and the test set. [4]

Table 2.1. Numeric example of underfitting and overfitting cases

	Case of underfitting	Case of overfitting	Case of good fitting
Human level	0.1	0.1	0.1
Training set level	0.8	0.15	0.16
Test set level	0.85	0.6	0.17

In order to conclude if our trained model is correctly trained, we must see carefully the error obtained in both the train set and in the test set. As we can see in table 2.1 shows a representative example of underfitting and overfitting and good fit cases. As we can see in the first column the training error (bias) is high we can say that is underfitting and we conclude that has not yet converged on an optimal minimum. In the second column, it's overfitting the error of the train set and the test set has high variance but the error bias is reduced. In the third column, we see a reduced bias error but the variance error is also reduced, we would be faced with a well-trained model. In order to face the cases of overfitting or underfitting, there are a series of strategies that allow to minimize them: [4]

- In the case of underfitting. The Possible strategies are to develop use a deeper network, improving its configuration, use a more advanced optimization algorithm, or change the whole structure of the network.

- In the case of overfitting. Most of the time it happens because of the small data set size [27], so it is reduced by increasing the number of samples in the train set or applying regularization techniques.

2.4.5. Deep Neural Network: Optimization

As we have seen in the previous subsection, optimization configuration is very important so choosing the correct optimization configuration for the network will help us to reduce the different predictions and correct values obtained. In this subsection, we will see the options we find when optimizing a neural network.

A. Gradient descent optimizer. There are many optimizers like Adam, SGD, Adagrad, RMSprop and so on the main function of the optimizer is to update the weights of the nodes in order to minimize the error of the cost function. We pay special attention to the Adam optimizer which combines the advantages of two optimization methods that are AdaGrad and RMSProp, obtaining a robust optimizer that adapts to a large number of problems [28].

B. Variants of Gradient Descent. Is an algorithm that calculates the error for each example in the training dataset and after all the training examples evaluated then it updates the model One cycle or one iteration through the entire training network dataset is called a training epoch. Depending on the amount of data used before updating the weights, we find 3 variants of gradient descent:

• **Batch Gradient Descent.** The cost function is not applied until all samples of the dataset have been passed, calculating the average error of the whole dataset to update each node by the backpropagation method.

• **Stochastic Gradient Descent.** With each new sample, the weights will be updated.

• **Mini-batch Gradient descent.** It is an intermediate point between the two previous ones. First, the dataset is subdivided into mini-batches. Through the hyperparameter "batch-size" we configure how many samples will compose a mini-batch and the descending gradient method will calculate the average error and will update the weights when the samples that compose each mini-batch are passed.

The most common method is the mini-batch gradient descent. This method combines the robustness of stochastic gradient descent and the efficiency of batch gradient descent. The "batch-size" is a parameter that is configured.

C. Learning rate. It is one of the hyperparameters and it can affect the training process. A learning rate too small gives rise to very slow convergence, while a learning rate that is too high cause divergence. In addition, the optimal learning rate is not a constant value throughout the entire training process. It is important to start with a sufficiently high learning rate and reduce it as the model converges to the global optimum. By implementing a learning rate scheduler we can regulate and modify the learning value. The general functioning of the various learning rate scheduler is based on the fact that after each epoch an action policy is applied in which it is defined in what circumstances [4].

It is necessary or not to reduce the value of the learning rate and in what proportion it is reduced.

D. Weight Initialization Techniques - Transfer Learning. When building a neural network the nodes, a previous step to train the model is to initialize the weights of the nodes. There are different options ranging from Random initialization to more sophisticated methods such as applying the He [11] or Xavier [29] initialization techniques. Another widely used technique with satisfactory results is called Transfer Learning [30]. Transfer learning consists of reusing the weights of a previously trained model against a very large dataset such as the Imagenet dataset [31] and with those reused weights start the training process being able to distinguish certain forms and characteristics inherited from the previous training in the early stages of training.

2.4.6. Deep Neural Network: Regularization

To avoid the result of overfitting we have several techniques to deal with it. Let see some.

A. **Data Augmentation**. The main concept of training a model is that when the training data set is bigger the better performance we will get. But sometimes the data set is limited and it is difficult to obtain new sample data. So in order to increase the size of the data set, we can apply a certain transformation to the training data set such as zooming, moving the image, cropping, applying some kind of distortion or noise in this way we can increase the number of the training data set.

B. **Dropout or Dropblock**. Dropout [32] is a regularization technique in which in the training process a number of nodes are randomly canceled or ignored with each new sample. The ultimate goal is to reduce the interdependent learning amongst the neurons by facilitating generalization in the training process. Dropout is an effective regularization method and provides important improvements to the training process, but its use is limited to fully connected layers. In convolutional networks where information is spatially correlated, although we apply dropout information continues to flow resulting in overfitting without dropout having any beneficial effect. For convolutional networks, there is a dropout variant called Dropblock [33]. Dropblock cancels blocks of nodes to avoid the propagation of some features maps towards the next layers and it achieves its function of regularization.

C. **Early stopping.** Early stopping means in the training process monitoring the certain matric and if the monitoring detects that the model is not improving or on the contrary, if the variance increases at that moment the training process will automatically stop.

2.4.7. Convolutional neural network

A convolutional neural network is like the human visual cortex when the brain process an image. CNN is more specialized in image processing and classification. Each node covers a region of the image and together with the rest of the nodes the whole image is covered. In CNN layers are combined and different layers. [4] The image is a multidimensional input where the pixels that compose it maintain a relation of local correlation with the neighboring pixels if we used feedforward neural network the number of necessary parameters would make its training unfeasible. Therefore, to work with images arise convolutional neural networks [34].

- **Input Layer:** this layer accepts raw image data and forwards it to the next layer for more extraction of the features.
- **Convolutional layer**. This is where the features extracted from the input image data. It will apply different filters or kernels, which contain the trainable weights, and which run through the entire image executing convolution operations and generating a feature map. Each and one of the filters will be specializing in detecting certain characteristics. In the first layers, they detect simple patterns such as different lines or curves and in the later layers, they detect more complex characteristics or patterns such as objects or shapes. After a convolutional layer an activation function is applied, most commonly using the rectified linear unit (ReLU). [4] the k^{th} output feature map y_k can be computed as:

$$y_k = f(w_k * x) \tag{2.1}$$

- **ReLU (Rectified-Linear Unit):** After the convolution layer the next layer is the Rectified Linear Unit or.ReLU. For fast and effective training this layer replaces the negative-number of the convolution layer with zero (0).
- **Pooling layer**. This layer has the function of reducing the dimensionality of features maps. The most common operations are on the one hand to detect the maximum value of the subsample region discarding the rest to this operation or to calculate the average of the elements of each region [35].
- **Fully-connected layer**. This is one or more fully connected layers and is located close to the output. These layers learn the relationships between features of maps and obtain output predictions that minimize the cost function.
- **Softmax Layer:** we find this layer just before the output layer and this layer for each class this layer will give the decimal probability. The decimal probabilities are between 0 and 1.

As we can see in Figure 10 the different layers are combined following a logical sequence. The convolutional layer and pooling layer take care of extracting the features maps from the input image. And in the last layers are located the fully-connected layers that carry out the classification function, reaching a prediction from the feature map extracted from the image.

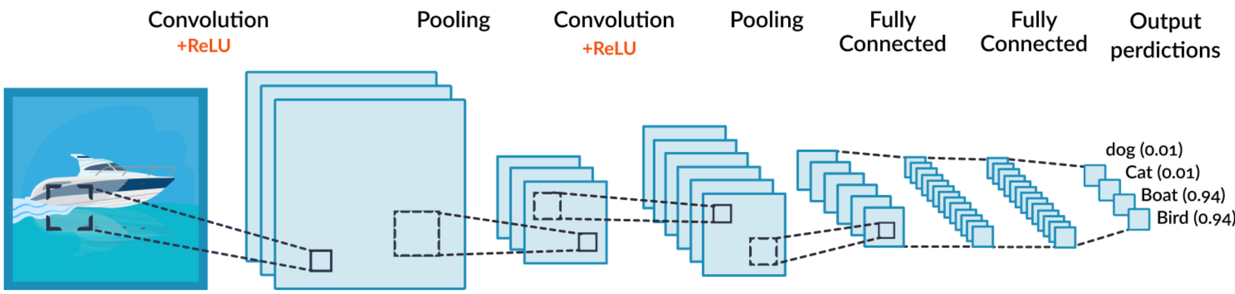


Figure 2.10. Basic structure of a neural convolutional network

2.5. Transfer Learning

As we have seen earlier one component of the machine learning algorithm is the description of the task to be solved, which is learned by the algorithm in the learning process.in deep learning, models can be trained with a large number of data and learn the model weight and bias during the training. And these weights and biases will be transferred to other network models for testing. The new network model can start with pre-trained weights [57]. A pre-trained model is already trained in the same domains. There are many pre-trained architectures are available, and why we use this pre-trained model are mentioned below as:

First, its main reason is that to train huge models and a huge number of datasets it requires more computational power.

Second, too much time taking to train the network up to a number of weeks. We can speed up the learning process by Training the new network with pre-trained weights.

Computer vision and natural language processing tasks use Pre-trained models as a starting point because the development of neural network models on this takes needs a lot of computing power and time resources. So for this complex and difficult task Pre-trained models are suitable for this purpose. These pre-trained models are trained with different data sets with sufficient information so that they can perform good and well with simpler tasks. In this way, the pre-training output weight can be used for initialization. Instead of choosing a random starting point. For basic shapes, images, and structures like spots and lines can be recognized in the early kernel of the network, and the detection of these basic features is very important to improve the performance of the network because the model doesn't have to learn all the features from the scratch. Depending on the tasks of the models or the example of the network of pre-trained networks like ImageNet, wordNet, and MNIT whether to classify images or words or numbers.[34]. For image classification The pre-training network usually trained on the very large amount of image database IMAGENET, which contains millions of different images depending on different objects, such as different mammals, cars, flowers, medical images, etc.[6] [31].

ImageNet organizes prepare a competition every year. The challenge is about image classification and some participants of the challenge propose there network architecture and the most successful are VGGNet, AlexNet, GoogleNet and ResNet [36], [37]. Architecture can be used as a pre-training model for different types of classification. Fine-tuning means Continuing to train new data using pre-trained network weights, which usually very good for training CNN, because of the need for a huge number of training data and computing power [38]. Most of the time we applied fine-tuning when we dealing with small amounts or limited data sets, because, for sufficiently large data sets, this may not be necessary [39], [38]. We have to modify some layers of the architecture In order to make the pre-trained model compatible with the new tasks of the network. It is beneficial to train more layers if the task of CNN is to classify images until the performance of CNN is improved, the fine-tuning technique is to gradually add layers [38].

2.5.1. Recently Proposed Deep Learning Architectures

Computer scientist Jan Lecken in the late 1990s proposed that CNN is used for image classification and recognition because of its high accuracy inspired by human visual recognition [34], [40]. A convolutional neural network (CNN) is a complex feedforward neural network. CNN follows a hierarchical model. It builds arrange like a funnel. At last, it gives a completely connected layer, where all the neurons are connected to each other and the yield is prepared. There are a few CNN designs with diverse numbers of CONV, pooling, ReLU, and fully connected layers. A few of the foremost common CNN structures are:-

- **LeNet:** In the 1990s, Yann Lecun, Leon Bottou, Yosuha Bengio, and Patrick Haffner proposed a neural network structure they called lenet-5 for handwritten and machine-printed character recognition [41]. LeNet is the first Convolutional architecture, which consists of two convolutional layers with ReLu and average pooling layers, followed by another convolutional layer, which is used for flattening, then two-fully-connected layers and ultimately one softmax-layer. This architecture is simple and clear that’s why it is mainly used as the first step in the teaching of CNN.

- **AlexNet:** in 2012 designed by Alex Krizhevsky is a much deeper neural network than the LeNet and it is one of the common CNN. -(ReLU) Rectified- Linear Unit it speeds up the network by adding non-linearity. This network has five convolutional layers,-three fully connected layers followed by the output layer and also contains 62.3 million parameters.[6]and was submitted to the ImageNet Large Visual Recognition Competition (ILSVRC) and achieved remarkable results.
- **VGGNet:** in 2014 proposed by Karen Simonyan and Andrew Zitherman and it is one of the highest performance CNN [40]. The full form of VGG is Visual Geometry Group. Normally VGG network contains VGG16 and VGG19. In this model the large size kernels are replaced with the multiple numbers of 3x3 filters, because of this we extract complex features at low cost.
- **GoogleNet:** In 2014 it was proposed From Google by Szegedy et al.[36] GoogleNet achieves good accuracy but it required high computational power because the order of calculations is very high. GoogleNet was replaced with average pooling after the last convolutional layer instead of fully-connected layers at the end; this will reduce the number of parameters. In addition, it uses a pooling layer at the top of the convolution layer instead of a completely associated layer. The two models utilize other forms of the starting module. The thought of the initial module is that we will utilize all convolution filters at the same time, rather than choosing a convolution filter. In other words, it employments numerous convolution filters to handle particular inputs, as well as application pools. At last, the results of all these forms are linked together [36].
- **Inception**: Inception was for the first time proposed by Google as an advanced version of the famous vision model ‘inception’. It was presented in a conference on the Association for the Advancement of Artificial intelligence (AAAI) 2017 by Christian Szegedy and Sergey Ioffe and Vincent Vanhoucke and Alexander A [36].

2.6. Related Work

As we can see from the previous section. Deep Learning, especially in recent years, has become experienced a strong momentum and enormous attention. As a result of this different numerous studies have emerged that analyze the influence of Deep Learning and mainly convolutional neural networks in the field of health and imaging medical diagnosis [42, 43, 22, 44, 45, 43]. All of them reflect the enormous interest and possibilities offered by the application of Deep Learning in a medical image. Among these studies, we highlight the paper [22] in which a complete compilation of more than 300 works is developed, whose common nexus is the application of Deep Learning in medical imaging. In the detection and diagnosis of some of the pathologies with a greater concentration of affectation and mortality, we find several works based entirely in the implementation of neural networks: Treating cancer [46, 47], tumor [48, 49, 50] or cardiovascular problems [31, 32, 33]. If we focus on the pathology of the respiratory system that concerns us, the detection of cases of pneumonia, we find several works have incorporated the use of neural networks in their proposals for implementation [51, 52, 53, 54, 55, 56]. Within the previously mentioned works, we highlight a reference study in the diagnosis of cases of pneumonia on chest x-rays using Deep Learning called CheXNet [56]. In this work published by Stanford Machine Learning Group1, they implement a 121-layer Densely Connected Convolutional Neural Network and whose results will be used as a reference to contextualize our results in classifying pneumonia cases.

Convolutional Neural Networks (CNN) is one of the best and the most successful models in performing various tasks such as classification and object detection [4]. The advancement of the models included a theoretical study, iterative prototyping, and experimental evaluation of the output. The current network, 169 layers DenseNet, by Pranav et al. (2018) on the abnormality detection task, the execution was lower than the worst radiologist in 5 out of the 7 studies, and in general model, the execution was lower than the most excellent radiologist. We created the ensemble200 model which scored 0.66 Cohen Kappa which was lower than the DenseNet network (Pranav et al, 2018) but the model execution with the F1 score outflanks the DenseNet model and its Cohen Kappa score changeability with the different studies is lower as the most excellent cohen kappa score on the upper limit studies is 0.7408 (Wrist) and the least is (0.5844) hand. The ensemble200 model outflanked DenseNet demonstrates on the finger studies with a Cohen Kappa score of 0.653 appearing decreased execution inconstancy on the model performance.

The authors utilized a gradient-based ROI localization calculation to identify and spatially locate pneumonia in CXRs. They released the largest collection of the National Institutes of Health (NIH) CXR dataset that contains 112,120 frontal CXRs, the associated labels are text-mined from radiological reports using natural language processing tools. The authors reported an AUC of 0.633 toward detecting the disease. The authors of [56] used a gradient-based visualization method to localize the ROI with heat maps toward pneumonia detection. They used a 121-layer densely connected neural network toward estimating the disease probability and obtained an AUC of 0.768 toward detecting pneumonia. The authors of [58] used an attention-guided mask inference algorithm to locate salient image regions that stand indicative of pneumonia. The features of local and global network branches in the proposed model are concatenated to estimate the probability of the disease. An AUC of 0.776 is reported for pneumonia detection.

Chapter Three

3. Data and Methodology

As the aim of this project was to develop an algorithm to recognize pneumonia infectious disease in images. The training was conducted on a normal laptop windows 10, Core i5-6500U with 1.8GHz clock speed, 8GB of RAM, 64 bit with CNN was proposed in order to perform the binary classification. In this chapter, first, we will describe the datasets we used for our experiments and we will introduce all the details about the algorithm and data processing implementation. This thesis also introduces the implementation of two algorithms, framework, network structure, and network training.

3.1. Datasets

In this study the subset of the Chest X-ray images was selected from retrospective cohorts of pediatric patients from Guangzhou Women and Children's Medical Center, Guangzhou was used to train and validate our convolutional neural network (CNN) classifier. All chest X-ray imaging was performed as a portion of patients' schedule clinical care. For the investigation of chest x-ray images, all chest radiographs were at first screened for quality control by expelling all low quality or incoherent scans. At that point, the diagnoses for the images were then positioned by two medical master doctors before being cleared for preparing the AI framework. In order to account for any reviewing errors, the assessment set was moreover checked by a third expert. [8][9][10]

This is a huge dataset on the X-Ray images and several data packages have been released to the research community. In this paper's experiments, the authors obtained a total of 5,864 chest X-ray images from children, including 3883 characterized as depicting pneumonia and 1349 normal. The dimensions of images are varied ranging from 384 x 127 to 2772 x 2098. The percentage of the dataset in each serving has been arranged as follows:

• Train Set: 80%. • Validation Set: 10% • Test Set: 10%.

This subset included different subjects and different age levels who underwent the process of Radiography and chest x-ray imaging scans. Classification of pneumonia infectious disease and normal health images requires several steps, from preprocessing to classification, which leads to the development of end-to-end pipelines.

The chest X-ray images are arranged in 3 folders (Train, Test, Val) and each contains subfolders for each image class (Pneumonia/Normal). There are around 5,864 X-Ray images (JPEG) and 2 class categories (Pneumonia/Normal).

Table 3.1. Dataset and its characteristics

Category	Training samples	Test samples	File type
normal	1349	234	jpg
bacterial	2538	242	jpg
Viral	1345	148	jpg

Figure 3.1 shows the number of images in each of the categories for our datasets. Map_character s = {0: ‘No Pneumonia’, 1: ‘Yes Pneumonia’} as we can see the data from the table above all th e images with pneumonia is 3883 and normal or non-pneumonia images are 1349. As we mentio ned in the previous section, VGG-16 architecture is used as a CNN classifier in this study. we us e the TensorFlow Deep Learning Framework as backend to train our model.

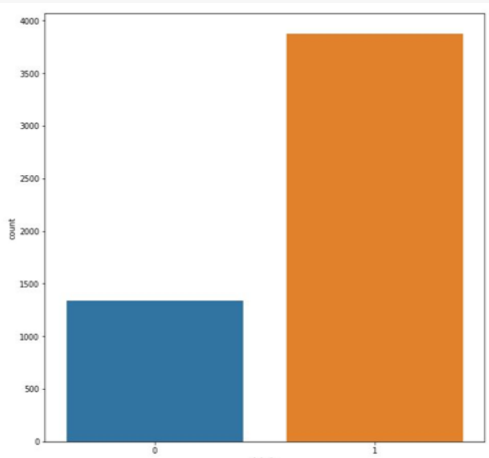


Figure 3.1. Illustration or graphical representation of our data

3.2. Image Preprocessing and Augmentation.

Deep learning needs a big number of data sets to obtain a good accuracy result. However, there may be difficult to get many datasets. Especially medical data sets to find a medical data set is a very expensive and time-consuming process. Data augmentation is a very good solution for this problem and it also helps us to avoid overfitting and helps to get high accuracy. Our data from Guangzhou Women and Children's Medical Center was limited. For this reason, data expansion and data augmentation have been implemented as a solution to the problem. Therefore, preprocessing such as extraction, smoothing, and normalization and converting the image. However, chest x-ray images are sensitive to excessive changes, because the characteristics of pneumonia must be preserved. To change the image using contrast enhancement or another color. For the image enhancement, it was limited to vertical and horizontal flipping, rotating, adding some noise and changing the brightness and contrast. In order to not lose the image originality, we applied slightly to change the brightness and contrast. For adding the noise and adjustment of the brightness we used python using the module skimage exposure. Skimage-util-random-noise model is used to add noise and also we added types of noises that are speckle noise, position noise, and Gauss noise. Before cropping the image for the rotation of image we have considered coordinates of cropping. We have applied a rotation matrix for each coordinate of the image and also for each corresponding pixel strength. A customized function is created in python for extending the image with the coordinates of clipping the image as input.

In this work, for the training time data augmentation method we used different augmentation methods such as height_shifting, width_shifting, zooming, shear range, flipping, and rotating at 50-degree angles. Transfer learning is also one of another performance-enhancing method in deep learning models, especially in CNNs. Transfer learning is the thought of overcoming the confined learning worldview and utilizing knowledge acquired for one task to solve similar tasks. These days some peoples prefer pre-trained CNNs from creating an entire CNN from scratch because it needs a large number of data sets.

There are three different transfer learning approaches in CNNs. These are feature extractor, fine-tuning and pretrained models [39].

We have implemented image preprocessing, or data augmentation methods to artificially increase the size and quality of the training dataset. This process will help to fix the overfitting problems and enhances the model's generalization ability increasing the accuracy during the training process. The settings deployed in image augmentation are shown below in Table 3. During the augmentation process, the rescale operation represents image reduction or magnification.

The first step is to rescale our data. Rescaling images is a common practice in deep learning because most images have RGB values ranging from 0–255. These values are too high for most models to handle, but by multiplying these values by 1/255, we can condense each RGB value to a value between 0–1. This is easier to process for our model.

Next, we have shear_range which will randomly clip the image angles in a counterclockwise direction apply shear mapping or shear transformations to the data. The value "0.2" is the shear intensity, or shear angle.

Zoom_range this is for randomly zooming in on the images set to "0.2".

Horizontal_flip is set to "True" we want to randomly flip half of the images in our dataset.

Rotation_range means the range in which the images were randomly rotated during training, i.e., 50 degrees.

Width_shift and height_shift the horizontal shift or the width shift of the horizontal translation of the image is 0.2 percent, and the height shift is the vertical translation of the images by 0.2 percent.

We will be using the ImageDataGenerator() class from Keras for our data augmentation. Data augmentation will help us to expand our dataset. The more training data and the more variety the better. With more training data and with some of the manipulated data, the problem of overfitting becomes less as our model has to generalize more.

Table 3.2. Settings for the image augmentation

Method	**Settings**
Rescale	1/255
Shear_range	0.2

Zoom_range	0.2
Horizontal flip	True
Rotation_range	50
Width_shift	0.2
Height_shift	0.2

Then we will go to the path of our test, train, and Val folders and generate batches of augmented data using flow_from_directory() from Keras.

The first argument must be the directory where the data pull from.

The second argument is the size of the image or the target size or the dimensions of the images after they are resized.

As we are working on binary classification and we have two classes the third argument is “class_mode”, which is set to “binary”. This will return 1D binary labels. This dataset calls for binary classification because there are only two classes.

After we finish processing our data now we can build our model and training it, then we test our model and getting our results in the form of our accuracy scores.

The chest x-ray images contain reigns that do not contribute to diagnosing the pneumonia infection other than the lung's reign. Because of this, the model may learn some unused features representations from the underlying data. Expert-delineated lung masks are used as models to register with the objective pediatric chest x-ray images as a reference set of patient chest x-rays. When displayed with an objective chest radiograph, the Bhattacharyya remove degree by the calculation to choose the foremost comparable models of the chest x-rays. By computing the modeling of the correspondence between the model chest x-rays and the objective chest x-ray with the image features representing and identifying similar locations by applying the SIFT-flow algorithm. For the objective chest radiograph, this map transformation applied to the model. The image is cropped to the lung boundary in the size of the box to include all the image pixels or the lung pixels that consist of the ROI for this task. The chest x-ray baseline data cropped bounding box are resampled to the appropriate pixel size dimensions and mean normalized to assist the model in faster convergence. The lung boundary for the sample pediatric chest x-ray is shown in figure 3.2

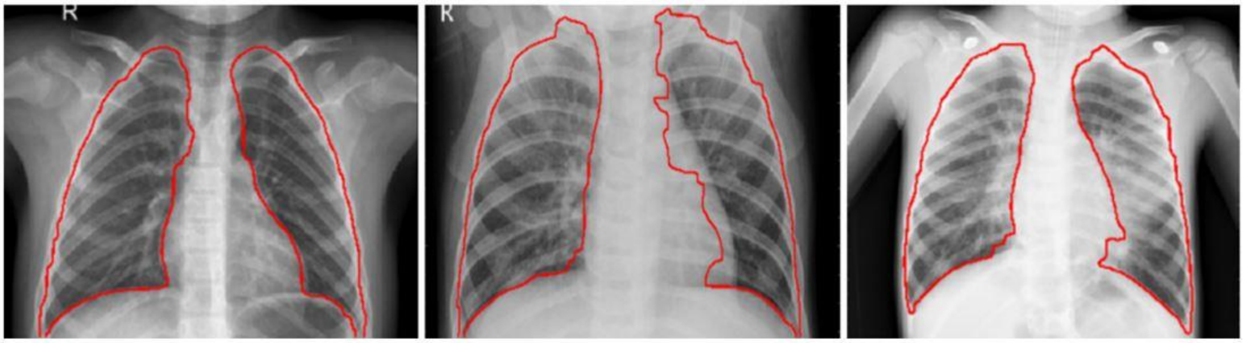


Figure 3.2. Detected boundaries in sample pediatric CXRs.

3.3. Methodologies

In this section, we will see studies preliminary on two deep learning models, VGG16 and Inception-V3, and explains why we choose these two models. The second part is the implementation of the algorithm, including the short explanation of the selected framework, network implementation, and network training.

Our experiments were conducted on a normal laptop windows 10, Core i5-6500U with 1.8GHz clock speed, 8GB of RAM, 64 bit by using anaconda on Jupiter notebook in python 3.7 on CPU system and create the CNN model based on Keras and Tensorflow libraries and we proposed two different models that are suitable with our libraries, such as VGG16 and InceptionV3. The reason that we chose these two models is that they have similarly succeeded in image classification issues. After each convolutional layer, multi-layer, automatic feature extraction optimization, normalization layer and so on improved classification, especially in the case of low variance between classes. In our experiment, the VGG16 model successfully has an accuracy of 94.3% for our first dataset and 85.6% for InceptionV3. The performance of our models was good in the classification process however the InceptionV3 model was more time consuming and it performs low accuracy. Therefore, for this experiment, we have chosen the VGG16 model to classify the images.

3.4. Deep Learning Architecture VGG16

VGG16 is proposed and trained by the Oxford's Visual Geometry Group (VGG) model proposed by K. Simonyan and A. Zisserman in their paper "Very deep convolution network for large-scale image recognition"[40]. The model scored first in ILSVRC image localization and second in image classification tasks. We customized the model architecture of VGG16 and evaluated its performance in the task of classification of chest x-ray images. This model originates from network architecture VGGNet, the number 16 represents the number of weight layers in the structure. The network consists of five convolutional layers and the size of the filter is 3 x 3. The largest pool layer separated five blocks and the three FC layers are added at the last. The prediction of the network as an output will be generated by the last layer of the FC layer. The number of channels corresponds to the number of classes in the prediction, and SoftMax is the activation function to return probability. The model achieves 92.7% accuracy in the first five bits of IMAGENET [31]. IMAGENET is a data set containing more than 14 million images belonging to 1000 classes. This is one of the well-known models submitted to ILSVRC-2014[6]. It improves Alexnet by replacing large-scale core size filters one by one (11 and 5 for the primary and second convolution layers, individually) with multiple 3*3-core-size filters. VGG16 gotten a few weeks of preparing and utilized Nvidia Titan Dark GPU. Before explaining the topology of VGGNet, we ought to describe more concepts utilized within the model. The following areas portray VGGNet features in more detail. Indeed in spite of the fact that ReLU activation functions don't require input normalization to dodge saturation, VGGNet topology uses local normalization after ReLU is connected within the conv layer. Indeed in spite of the fact that the ReLU actuation function does not require input normalization to avoid saturating zone, VGGNet topology uses a local normalization after using ReLU in CONV layers. They utilize exceptionally small (3 x 3) convolutional filters to extend the depth of the architecture to 16 and 19 layers. The VGG16 architecture consists of 12 convolution layers, some of which are followed by the largest pool layer, then four fully connected layers, and finally 1000-way SoftMax classifiers [40]. They named their findings VGG16 (Visual Geometry Group) and VGG19. These two models in 2014 submitted for the ImageNet Challenge and the team ranked first and second in localization and classification respectively.

3.4.1. Implementation of the network

TensorFlow was the selected framework and with the machine learning library Keras they used as the back end together. For this project, different models including VGG16, VGG19, and InceptionV3 are retrieved from the module application mentioned in the above section. In the early stage of the project, the Nvidia in-depth learning GPU training system Digits were also checked, and TensorFlow was used as the framework [61]. Digits provide an opportunity to try out networks such as LeNet, AlexNet, and GoogleNet, with pre-training on appropriate databases such as ImageNet [6]. Before the implementation of VGG19 network based on Tensorflow framework, a small test was done by using data sets from different networks. The trained network shows the decreasing of the loss and the improvement of the accuracy and it produces the most promising results. The selected pre-training network is VGG16 that is implemented from the pre-training weight from the database of ImageNet. We have modified the last layer of the network in order to use it for binary classification to classify images to process only two images. The class and activation function changed from SoftMax to Sigmoid. A sigmoid function is a mathematical function having a characteristic "S"-shaped curve or sigmoid curve. A common example of a sigmoid function is the logistic function shown in the below figure and defined by the formula:

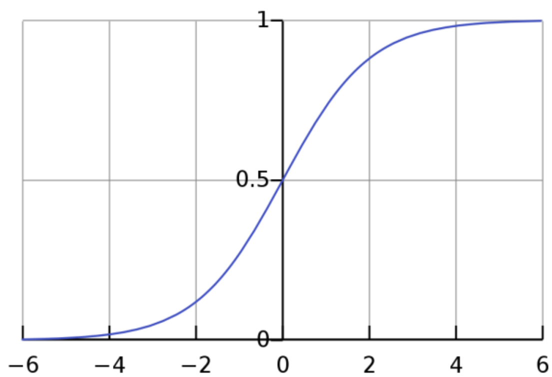


Figure 3.3. Sigmoid logistic curve

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}. \qquad (3.1)$$

To avoid over-fitting, two dropout layers have been added after the first and second FC layers. The final modified network architecture is shown in Figure 3.4.

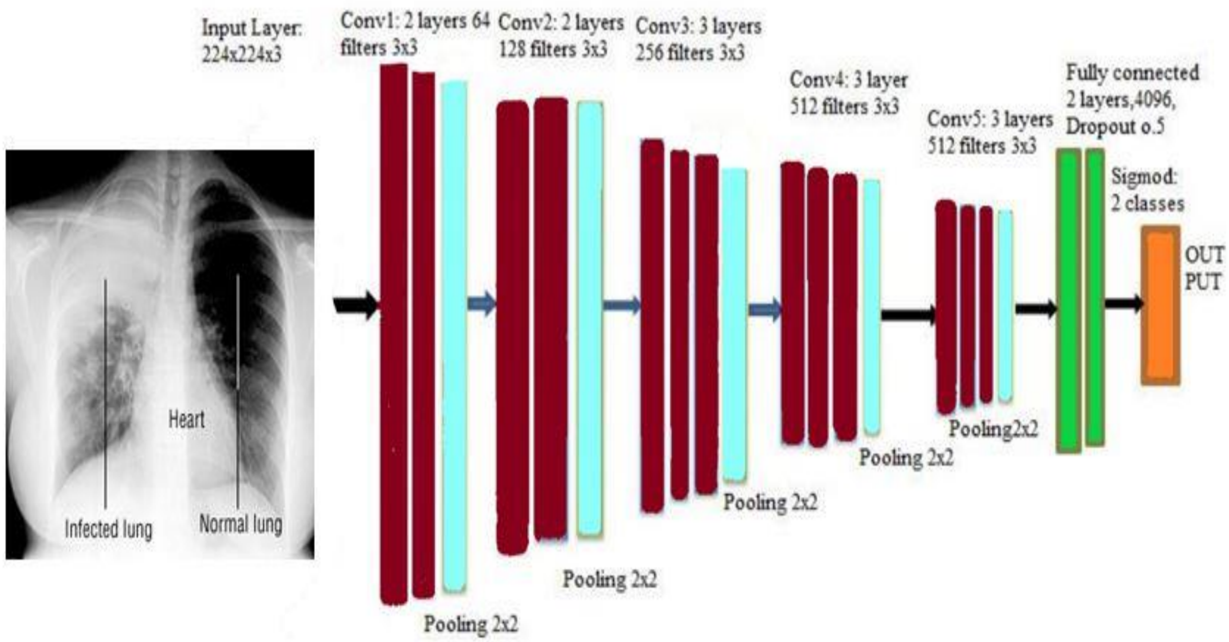


Figure 3.4. An illustration of the VGG16 architecture.

3.4.2. Training the Network

Finally in the network training phase to train the network using keras, the settings of the training needs to be compiled where the settings of the learning rate, optimizer and loss function need to be set. The same settings will be applied to pre-training using pneumonia infection chest x-ray images.

The input data to the convoulution1 layer is of a fixed size 224 x 224 image. Then the filters were used with 3x3 receptive field and the input image will pass through a stack of convolutional (conv.) layers, the receptive filed: 3x3 which is the smallest size to capture the notion of left/right, up/down, center. In one of the configurations, it also utilizes 1x1 convolution filters, which can be sometimes seen as the linear transformation of the input channels. The stride of the convolution is set to 1 pixel; the spatial padding of the convolution. Layers. Spatial pooling is carried out by five max-pooling layers, which follow some of the conv. Layers (not all the conv. layers are followed by max-pooling). Max-pooling is performed over a 2×2 pixel window, with stride 2.

The stack of convolution layers followed by three fully-connected layers with different depths in deferent architecture. The first two have 4096 channels each, the third performs 1000 way ILSRC classification and thus contains 1000 channels that is one for each class. The last layer is the sigmoid layer. The configuration of the fully connected layers is the same in all networks.

The super-parameter set before training includes time interval, learning rate decay, learning rate, loss function of the class weight, regularization parameter and dropout level. Each time when the network trained to evaluate the performance indicators the different levels of each hyperparameter are tested and adjusted several times before determining the most appropriate level. In addition to being pre-trained on ImageNet, the network was trained using a set of chest x-ray images from Guangzhou Women and Children's Medical Center before finding images of pneumonia infection diseases. The dataset utilized for the pretraining was not adjusted and the loss function was scaled once more with a weight factor comparing to the class indices to ensure the same significance for both classes amid the training. At the time of training first, we try to train with the number of Epoch 100 but according to time we decided to pre-training the model with 60 Epochs otherwise it would have taken too long before the test could be carried out. As also for the fine-tuning using the chest x-ray images the number of times was originally set too high, at 100 times, but we have adjusted according to the convergence of the measurements of the performance. Therefore, the number of times for fine-tuning was also finally set to 60. Using the generator by the time of retrieving images keras allows real-time processing of the images as they are collected from the correct location folder. The generator tool can also be extended with keras without having to save each image that is enhanced. However, since the images are cropped before being fed to the network, the keras provide the rotation produces artifacts such as lines and black edges in the chest x-ray images.

3.4.3. Overfitting in VGG16

As we mentioned before on section 2.4.4 overfitting means when the model learns the whole images in the training process, but not the features of the image. Most of the time overfitting happens when the model number of weight and parameters in the CNN model becomes high and when we have a small amount of the data set. Overfitting will cause degrade the performance of the model in the process of testing the datasets?

The number of parameters in the VGG16 model is 14 million. In order to avoid overfitting in this kind of model, we have to increase the number of our samples 10 times larger than the number of parameters. There are many ways to avoid the overfitting on CNN. The first method is to increase the number of our datasets or to reduce the complexity of the model, which is possible to do all the time. And the second method is to perform data augmentation or data preprocessing. VGGNet proposed or used to types of enhancement. The first to produce image translation and horizontal reflection [6]. The second one to form changes on the RGB intensity in the training images. And the last method to avoid overfitting is to increase regularization. VGGNet adds the Dropout method [55] to the model as regularization. Dropout has a specific probability (0.5 in VGGNet) that sets the output of each hidden neuron to zero and has. This means some of the neurons will be quiet and they will not move forward or backward simply means the model trains different networks or architectures for each input. However, all of these networks or architectures have weights. Reduce complex co-adaptation by reducing or eliminating one neuron's depending on the other neurons. This technique allows neurons to produce more features.

3.5. Inception V3 Architecture

Inception V3 is one of the transfer learning models which is used widely in image recognition and classification tasks. The inception V3 model attains very good accuracy on the ImageNet dataset. The model was a combination of many ideas that were developed by many researchers over the year. Based on the paper "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et al.

The model is made up of symmetric and asymmetric building pieces, counting convolutions, average pooling, max pooling, concats, dropouts, and completely connected layers. Batchnorm is utilized broadly all through the model and connected to activation inputs. Loss is computed through Softmax.

Inception V3 is a combination of a few symmetric and asymmetric building squares. These building pieces incorporate layers such as max pooling, normal pooling, dropouts, concats, completely associated layers where batchnorm is most broadly utilized. Inception mainly consists of two parts, the feature extraction part is carried out with a convolutional neural network and the classification part is dependent on the softmax layers. Inception V3 has pre-trained model which can easily classify 1000 classes which may have animals, flowers or be it non-living things. In the first part, the model only extracts the general features for the input images and the classification part is done in the second.

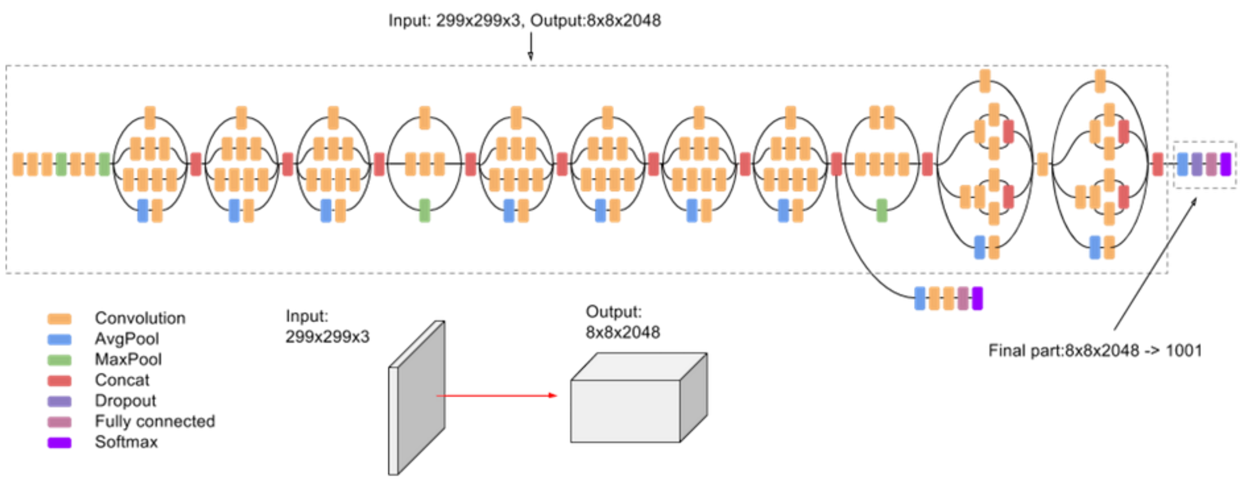


Figure 3.5. An illustration of inception V3 model

3.5.1. Preprocessing Stage

Image preprocessing is a very important part of the system and it can heavily influence the maximum accuracy that the model can give during the training. In the case of this model, the image size must be 299x299x3 pixels the images need to be decoded and resized to fit in the model.

However, simply just decoding and resizing the image's data is not enough to get very good accuracy. When the model first tested it was with a training data set of 1,281,167 images. Each pass over the set of the training phase we call it epoch. In order to improve the model image recognition capabilities during the training, it requires several passes. In the case of the model inception V3, the number of iteration or number of epochs needs to be somewhere between 140 to 200 range depending on the global batch size.

It is very useful to continuously alter the images before feeding them to the model and in this such a manner that a particular image is slightly different at every epochs. And also a well-designed preprocessing stage can significantly boost the recognition capability of a model. On the other hand, a too simple preprocessing stage can significantly cause an artificial ceiling on the maximum accuracy that the same model can attain during training.

Initiation V3 offers diverse alternatives for the preprocessing organize, extending from moderately straightforward and computationally cheap to decently complex and computationally costly.
Two particular flavors of such can be found in a file vgg_preprocessing.py and inception_preprocessing.py. File vgg_preprocessing.py characterizes a preprocessing stage that has been utilized effectively to train Resnet to 75% accuracy, but yields problematic results when connected on Inception v3.
File inception_preprocessing.py contains a multi-option preprocessing organize with distinctive levels of complexity that has been utilized effectively to prepare Inception v3 to accuracies within the 78.1-78.5% range when running on TPUs. This segment talks about the preprocessing pipeline. Preprocessing varies depending on whether the model is experiencing training or being utilized for inference/evaluation. At evaluation time, preprocessing is very clear: crop a central region of the image and after that resize it to the default 299x299 measure.

3.6. Model and solver definition

In this section, we will see how we have implemented our model in the Tensorflow framework and as a backend we use Keras. We adjusted the VGG-16 model epochs to 60 because when epoch 60 approaches to zero. Then we have applied the Tensorflow stochastic gradient descent (SGD) method is proposed to solve the problem. The SGD will generate training and testing networks for learning and evaluation purpose respectively. It takes the snapshots of the calculator state and the optimization process of the model. In order to learn the model, in each iteration process, the solver by calling the network forward it will calculate the output and loss and calculate the gradient by calling network backward. According to the solver method, the parameters will update in different ways. After updating the parameters, the solver will update its state according to the learning rate, method, and history. The changing of the intensity of the RGB channel was the second form of data enhancement in the training image. On the entire data set, we perform principal component analysis (PCA) on the set of RGB pixels. In every image training the multiple of the principal components we find proportional corresponding eigenvalues, which is multiplied with random variables extracted from the Gauss with an average zero and standard deviation of 0.1. The optimum value of the initializing learning rate is close to 0.01, then subtracts the constant value in every iteration. In every step of the iteration, the learning rate will be reduced by one gamma factor. Most of the time in deep learning the learning rate and momentum parameters are related.

In this work, the momentum parameter is _0.9. To make the model more stable and fast momentum smoothes the weight updating in the iteration process.

As we have seen previously, the VGG16 architecture was used as a CNN classifier in our study. We labeled our datasets into three folders. To evaluate and validate the effectiveness of our model, we have repeated our experiment six times for more than six hours. Parameter and hyperparameters were heavily turned to increase our model performance. And we have obtained different results but this study reports only the most valid one. As explained above, methods such as data augmentation, learning rate variation, and annealing were deployed to avoid overfitting problems. We have repeated our experiment six times on the jupyter notebook and on Kaggle online GPU system. Jupyter notebook is an open-source web application that can allow you to create and share documents containing codes, visualization, formulas and narrative texts. This notebook supports more than 40 programming languages including Python, C++, Scala, and Julia, etc. jupyter notebook includes data cleaning and conversation, statically modeling, numerical simulation, machine learning, etc. our result of this experiment will be discussed in the next chapter 4.

Chapter Four

4. Evaluation and Discussion

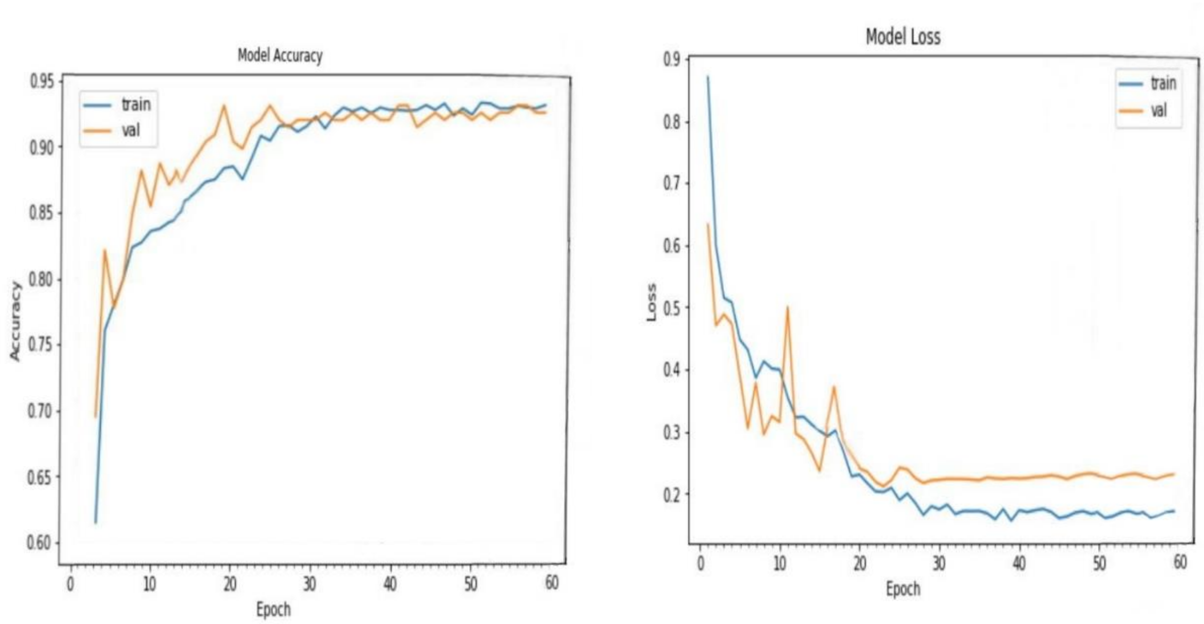
In this chapter, we will describe the result of our project, training strategies and test results. Our experiments were conducted on a normal laptop windows 10, Core i5-6500U with 1.8GHz clock speed, 8GB of RAM, 64 bit by using anaconda on Jupiter notebook in python 3.7 on CPU system and create the CNN model based on Keras and Tensorflow libraries and we proposed two different models that are suitable with our libraries, such as VGG16 and InceptionV3. We developed a model that can detect and classify pneumonia from chest x-ray images that are taken from the front view. The algorithm begins by transforming chest X-ray images into sizes that are suitable for the models. The next step is classifying and identifying the images by the convolutional neural network, which extracts important features from the input data images and classifies them. Before we train our model all the images dataset are resized for the target network model. Because the inceptionV3 model network as we have discussed in section 3.5 accepts images with a size of 299x299x3 dimensions whereas the VGG16 network accepts the image dimension size of 224x224x3. And also, all images pixels in 0.1 normalized range. We have trained these two models with the same parameters. But with a different number of epoch size of, binary_crossentropy selected as loss function, adam used as an optimizer, learning rate set as 1e-4, weight decay set as 0.9 and batch size set as 16. In order to avoid overfitting different strategies were used firstly, batch normalization was used in every layer. Secondly with a rate of 0.5 dropout method was used after fully connected layers. Lastly, the data augmentation was used in order to avoid the overfitting problem. Random datasets are labeled as binary classification, 80% of the images are allocated to training data sets, and the remaining 20% is used for testing. During the data conversion process, a total of 5,864 images were produced, including 3883 pneumonia infection disease and 1349 normal samples.

4.1. Accuracy of our models

Overall, the experiment results signify the robustness and effectiveness transfer learning convolution in the task of pneumonia detection. A highly good accuracy score is achieved by the two models the VGG16 model scores an accuracy of 94.3% and the lowest score accomplishing by the inceptionv3 model 85.6%. The results of this experiment are the accuracy of each type of prediction. In the process of classification, VGG-16 model and the inceptionV3 model were trained and tested with a large amount of data acquired from radiography chest x-ray images. In the model, a set of learnable filters are applied to extract low-level to high-level features from images. We work our experiments on the Jupyter notebook Python open-source library, with an average accuracy of 94.3%. Table 4.1 shows the accuracy of our two models for each experiment. As shown in the figure, we have achieved very good accuracy in all experiments.

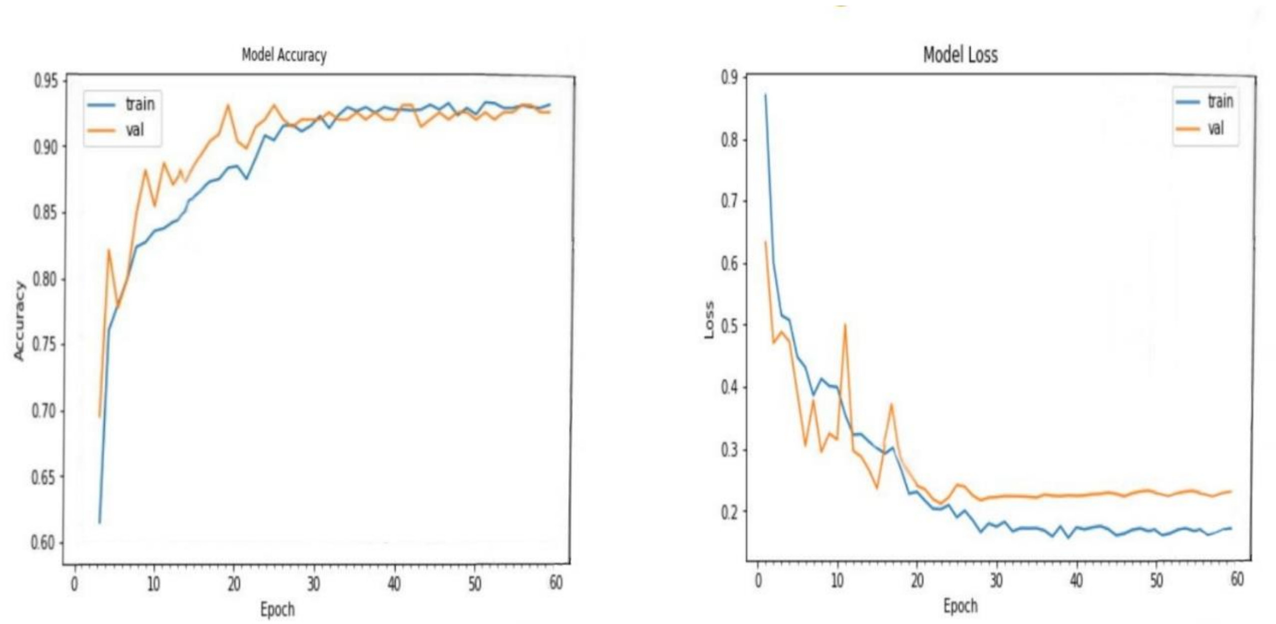
Table 4.1. The accuracy of our models in each experiment and the average accuracy of all six experiments.

Fold No	00	01	02	03	04	05	Avg
Accuracy of VGG16	94.59%	93.04%	95.12%	93.12%	93.3%	93.4%	94.3%
Accuracy of inception V3	85.64%	82.34%	84.44%	83.32%	86.33%	85.22%	85.6%



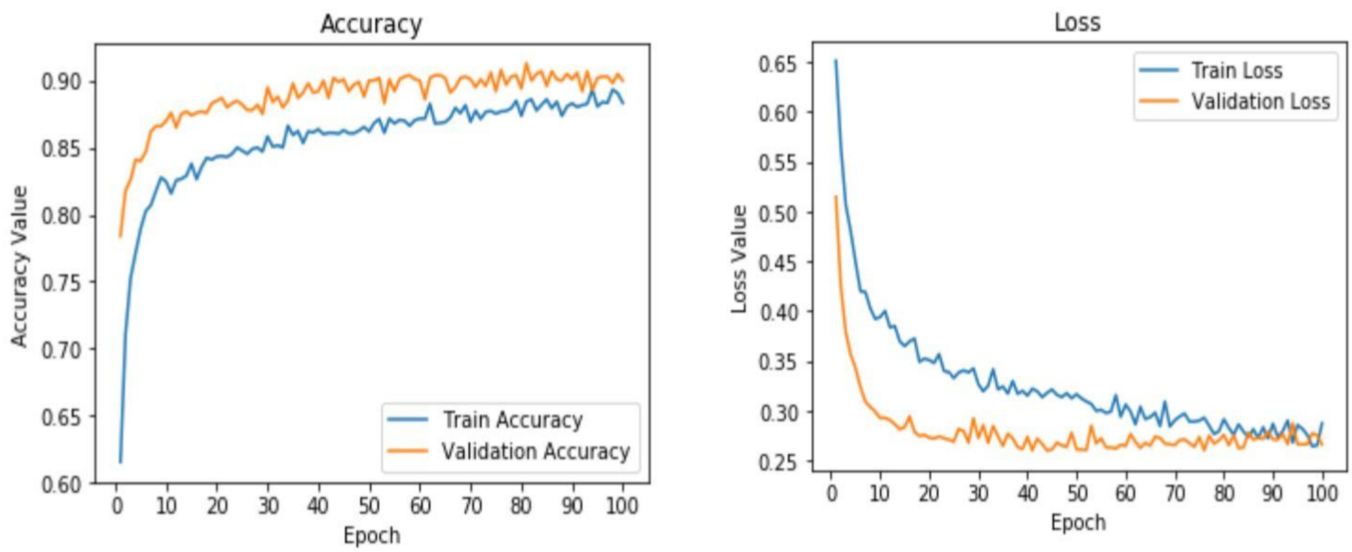
(a) (b)

Figure 4.1. (a) Accuracy and (b) loss graphics of the VGG16 model the fifth experiment in 60 epochs.



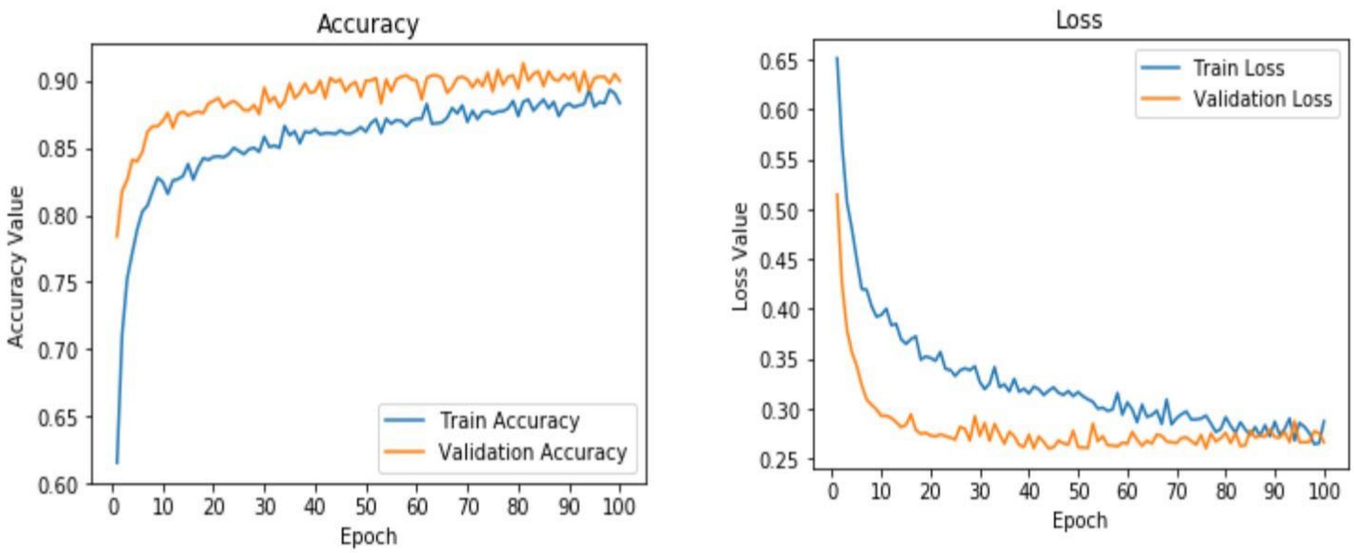
(c) (d)

Figure 4.2. **(c)** Accuracy and (d) loss graphics of the VGG16 model the last sixth experiment in 60 epochs.



(e) (f)

Figure 4.3. (e) Accuracy and (f) loss graphics of the inceptionV3 model the last or the fifth experiment in 100 epochs.



(g) **(h)**

Figure 4.4. (g) Accuracy and (h) loss graphics of the inceptionV3 model the last or the sixth experiment in 100 epochs.

Fig.4.1 fig 4.2 fig 4.3 and Fig. 4.4 shows the VGG16 and the inceptionV3 model's network's accuracy and loss graphics respectively. The results show that CNN architecture VGG16 is a more powerful model than the inceptionV3 model to classify our chest x-ray datasets. As we have seen from our project we try to compare two transferring models to classify and predict pneumonia infection disease based on chest x-ray images and based on our result we find VGG16 model is more powerful and accurate more than the inceptionV3 model. In order to better understand the model and the convolution layer, we visualized the weight filters and counted their information in every layer of the model for many samples of our test dataset. The proposed networks are implemented by using deep learning model Keras in Python programing language on a Windows operating system. The calculation run times of the Training and test of two models are given in Table 4.2. Given test time represents the evaluation time of 625 test images. The estimated time per image is 0.016 and 0.024 seconds Vgg16 and inceptionV3 models respectively. These results show that the models are suitable for real-time predictions.

Table 4.2. Network's train and test calculation time

models	Training time	Test time
VGG16	120 minutes	10 seconds
inceptionV3	200 minutes	15 seconds

4.2. Evaluation metrics

The confusion matrices and AUC achieved with the customized VGG16 and the InceptionV3 model are shown in Figures 4.5 and 4.6. We have observed that compared to accuracy the training metrics are poor. This is because of some noisy images are included in the training data to reduce bias, overfitting and to improve generalization of the model.

We have evaluated the two models by using 624 frontal chest X-ray images. The test dataset contains 234 normal images and 390 pneumonia images. The two models are evaluated by using the performance metrics such as accuracy, specificity, sensitivity, precision, recall, and f1 score. We can see the formulas below:

$$\text{Accuracy} = (TP+TN)/(TP+FN+TN+FP) \tag{4.1}$$

$$\text{Sensitivity} = TP/(TP+FN) \tag{4.2}$$

$$\text{Specificity} = TN/(TN+FP) \tag{4.3}$$

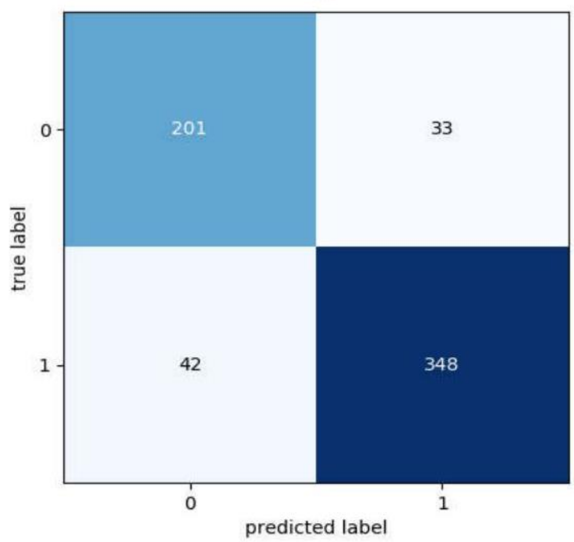
$$\text{Precision} = TP/(TP+FP) \tag{4.4}$$

$$\text{Recall} = TP/(TP+FN) \tag{4.5}$$

$$\text{F1 score} = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall}) \tag{4.6}$$

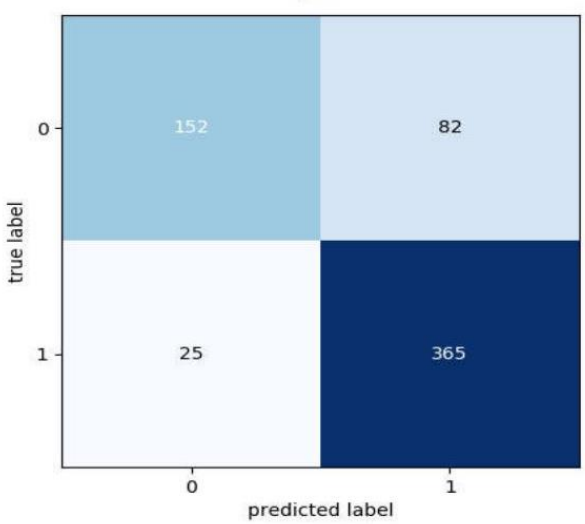
TP = represents the number of true positive, TN = represents the number of true negative, FN = false negative, FP = false positive respectively.

Across the classification task, we have observed that our model outperforms the current literature in all performance metrics. The customized VGG16 model demonstrates higher values for recall in classifying normal and pneumonia. VGG16 model under the two-class normal and pneumonia the TP, TN, FP and FN rates of 10 thresholds (from 0 to 1) after we give the image with pneumonia category we have got the probability that the image is classified as pneumonia category member is 0.9412, and the probability that the image is classified as other category was NC_0.0072 Pneumonia_0.00012. We used this data to generate a confusion matrix for 624 test images.



VGG16

Figure 4.5. The confusion matrix for testing 624 images in VGG16 our trained model



Inception V3

Figure 4.6. The confusion matrix for testing 624 images in inceptionV3 our trained model

We also calculated the ROC curve. The confusion matrix and accuracy calculated by the equations below. For pneumonia class value is 0.9459; for normal class is 0.9596. This shows that our result is very consistent. The performance of the classifier has been calculated with the true positive rate (TPR) AND false detection rate (FDR) as we can see below:

$$\text{Sensitivity True positive rate (T PR)} = TP/(TP+FN) \tag{4.7}$$

$$\text{Specificity True negative rate (TNR)} = \text{TN./(FP +TN)} \quad (4.8)$$

Among them, TP is a true positive (pneumonia classified as pneumonia), FN is a false negative (pneumonia misclassified as normal), TN is true negative (normally classified as normal), and FP is false positive (normally classified as pneumonia). In the field of medical false negative or FN and false-positive or FP is very critical and it must be low as much as possible. In another word, it means we have to have high TPR, F score, and accuracy, and at the same time, FDR must below.

4.3. Result Discussion and some limitations

The primary goal of our study was using a transferring learning approach to correctly diagnose pneumonia among normal chest X-ray images. Although the results we have found were very good there are some limitations in our model that we have to keep in consideration. The first and biggest limitation was that there was no history of the patients associated with the data in our evaluation model. Secondly, our data set was only frontal chest X-rays that we use to train the model but lateral views also helpful in the diagnosis of pneumonia. Thirdly, since the model exercises a lot of convolutional layers the model needs very high CPU and GPU powers otherwise it will take too much time in the process of computation. And another and very important thing was the size of the data set, as we know for the reliable and more accurate results we need a huge number of data sets. During this study, our main concern was limited data sets which are undoubtedly will affect the results because to train CNN we need large size of data sets it is more easier to achieve a good result and better performance with larger datasets. And the other challenge was the time span of the master’s thesis it was very limited.

Another possible advancement includes network pre-training. Pre-training utilizing images portraying different melanomas appears promising. This specified within the visualization section particularly by decreasing the loss of validation and training data and improving the accuracy, accuracy, and recall of training data. This shows that if proceeded, the pre-training may accomplish higher accuracy and lower losses. These pre-training results can be progressed by performing extra hyper parametric alterations. However, the time required to complete the desired number of epochs is unreasonably long, somewhat since of the huge amount of images, so no extra testing is carried out. If this is often overcome, extra data sets can be utilized for pre-training, so that the network can train more data. One possibility is to train the network on images with higher likenesses than chest x-ray images, counting other digital photos. One of the issues with network execution is the shortcomings of KERA. The library has numerous valuable tools and is simple to utilize, but the structure of building a network does not permit altering a single building block in numerous cases. In spite of the issues within the implementation process, there's still hope. But we should be noted that the results are the same regardless of the setting of the dropout level.

4.4. Comparison between Our Experiment and Previous Works

As shown in Table 4.3, many studies in the past utilized a binary classification between two categories of pneumonia and normal control subjects. And also many studies in the past used both multimodal methods and convolutional neural organize models, and most of them achieved remarkable results about in comparison with pneumonia and non-pneumonia. However, when we compare our experiment with their experiment our experiments and models are more satisfactory than their results. In our experiment, we applied an existing CNN transfer learning the existing deep learning model, compared with another method we have significantly improved compared with other methods to classify pneumonia infection. The table below will show us the summarization comparison between our experiment and other experiments based on their model and accuracy reported in each reference. Chest x-ray datasets and different techniques were used to classify pneumonia and normal.

As shown in the Table, there are different studies that are made in the past and the studies use different methodologies and use different strategies and all the studies use chest x-ray data sets for their work as we can see our experiment has more accuracy from previous works. We have made researches tried two to see all the previous studies and based on our research the studies that we found we tried to mention them on the table below.

Table 4.3. Based on the accuracy of the model, our results are compared with previous studies.

reference	Dataset	Method	accuracy
[62] Ayan Enes,	Chest x-ray images	VGG16 & Xception	0.87% , 0.82%
[63] Asnaoui, Khalid E	Chest x-ray images	VGG16 and VGG19	86.26 % , 85.94 %

[64] Blaku, Flokart,	Chest x-ray images	VGG16, VGG19 & inceptionv3	0.806%, 0.758% & 0.673
[65] Liu, Mao,	Chest x-ray images	CNN	90.25%
Our model	Chest x-ray images	VGG16 & inceptionV3	94.3% &85.6%

Chapter Five

5. Conclusion and Future Work

5.1. Conclusion

From the beginning of this master's thesis, our main goal was to develop a model or to develop an implementation that would allow us to accurately diagnose cases of pneumonia from chest x-ray images that can provide a binary classification of the presence or the absence of pneumonia by using transfer learning. In this framework, we adopted the transfer learning approach and used the pertained architectures VGG16 and Inception V3 trained on the ImageNet dataset, to extract features.

Our studies support the idea that deep learning strategies can be utilized to simplify the diagnostic process and improve disease management. Whereas pneumonia diagnoses are commonly confirmed by a single specialist, allowing for the possibility of error, deep learning strategies can be respected as a two-way confirmation system. In this case, the decision support framework gives a diagnosis based on chest X-ray images, which can then be confirmed by the attending doctor, radically minimizing both human and computer mistakes. Our results propose that deep learning methods can be utilized to move forward to diagnosis relative to traditional strategies, which may improve the quality of treatment. When compared with the past traditional strategies, our approach can viably distinguish the inflammatory region in chest X-ray images.

This paper fundamentally aims to progress the medical adeptness in areas where the availability of radiotherapists is still constrained. Our study encourages the early diagnosis of Pneumonia to prevent unfavorable results (including death) in such remote regions. So far, not much work has been contributed to particularly to detect Pneumonia from the mentioned dataset. The advancement of algorithms in this domain can be profoundly advantageous for giving better health-care services. In this Classifier. We observed the execution of two pre-trained CNN models. Within the process of classification, the features from low to high levels are learned, with an average accuracy of 94.3%. This is a prominent result, compared with other strategies utilized in past pneumonia infection classification.

As a conclusion, we will implement image detection for pneumonia detection by using an artificial neural network. With a limited number of medical staff, an automated system can significantly decrease the tedious manual labor involved in diagnosing large quantities of pneumonia x-ray images. The algorithm can be considered successful in many views, and the results prove the importance of the method. However, some improvements can be made in implementation and further development should be considered.

5.2. Future work

Before the execution and implementation of deep learning, there are still a few enhancements that must be made. This is a solution based on the field of the medical sector. The results are promising, but as we specified earlier, adding datasets can create better performance in classifying pneumonia infection disease. Larger data sets also introduce the possibility of including more classes to categorize the other stage of pneumonia infection. Future developments also include the addition of extra non-imaging parameters, which are usually assessed in determining the diagnosis of pneumonia. This presents the issue of different sorts of input data, but if this issue is illuminated, a complete automatic pneumonia diagnosis framework can be created. In any case of the complexity, the possibility of creating a variety of accessible tools through mechanized categorization emerges. This solution can be created as a Smartphone Application and used as a support framework in medical care. Medical staff can scan the lungs in seconds and get predictions. If this target group is taken into consideration, the application can offer assistance to medical staff work and reduce their workload. Within the long run, this will offer assistance patients and who will provide accurate treatment, faster, less discomfort and better quality. At last, considering the limited data set provided and the promising results gotten, a deep learning-based image classification solution for chest x-ray images ought to be considered reasonable. The high accuracy, recall and accuracy values of the verification set show that the potential of machine learning is based on solutions when comparing different settings. As a conclusion, the proposed solution may be a great beginning point to assist health care laborers to diagnose pneumonia disease. It includes money related to viewpoints and patient welfare.

Reference

[1]. Valiant,L.,G.A theory of the learnable commn,ACM, 27(11): 1134-1142,Nov 1984. 1. Valiant,L.,G.A theory of the learnable commn,ACM, 27(11): 1134-1142,Nov 1984.

[2]. K.Elaiyaraja & M.Senthil Kumar, Fusion Imaging in Pixel Level Image Processing Technique-A Literature Review, 2018

[3]. R. Hopstaken, T.Witbraad, J. Engelshoven, and G.JDinant, "Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory infections," in Clinical Radiology. The Royal College of Radiologist, 2004, vol. 59, pp. 743–752.

[4]. Vázquez Enríquez, Manuel. *A deep learning approach for pneumonia detection in chest X-ray*. Diss. 2019.

[5]. Singh, "Convolutional Neural Networks for Image Classification." [Online]. Available: https://www.completegate.com/2017022864/blog/deep-machine-learning-images-lenetalexnet-cnn. [Accessed: 26-Apr-2019].

[6]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, USA, 2012, pp. 1097–1105.

[7]. B. J. Erickson, P. Korfiatis, Z. Akkus, and T. Kline, "Machine learning for medical imaging," Radiographics : a review publication of the Radiological Society of North America, Inc, vol. 37, no. 2, pp. 505–515, Mar. 2017.

[8]. https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

[9]. http://dx.doi.org/10.17632/rscbjbr9sj.2#file-41d542e7-7f91-47f6-9ff2-dd8e5a5a7861

[10]. Kermany, Daniel;Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography(OCT) and Chest X-Ray Images for Classification", Mendeley Data, v2http://dx.doi.org/10.17632/rscbjbr9sj.2

[11]. "Revised who classification and treatment of pneumonia in children at health facilities: Evidence summaries. "World Health Organization, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK264164

[12]. https://www.who.int/news-room/fact-sheets/detail/pneumonia

[13]. https://www.medicinenet.com/pneumonia_facts/article.htm

[14]. Source: National Heart, Lung, and Blood Institute; National Institutes of Health; U.S. Department of Health and Human Services.

[15]. Asthma UK and British Lung Foundation Partnership is a company limited by guarantee 01863614 (England and Wales).

[16]. Hippocrates. The genuine works of Hippocrates, 1:324. In: Translated from the Greek with a preliminary discourse and annotations by Francis Adams. London: Sydenham Society; 1849.

[17]. Laennec RT. A treatise on the diseases of the chest and on mediate auscultation. In: Translated from the 3rd French ed. by J. Forbes, with notes of Prof. Andral from the 4th ed. New York: SS & Wm. Wood; 1838.

[18]. Rokitansky C. Inflammations of the lungs (pneumoniae). In: Manual of Pathological Anatomy. London: Sydenham Society; 1852.

[19]. Wells EF. Introduction to study of pneumonic fever. III. Nature. JAMA 1889; •••:332. http://dx.doi.org/10.1001/jama.1889.02400870008001a.

[20]. United States Bureau of the Census. Manual of the international list of causes of death: based on the fourth decenial revision by the International Commission, Paris, October 16 to 19, 1929. Washington: Government Printing Office; 1931.

[21]. ICD-10 Version. 2016. World Health Organization website; Available from: http://apps.who.int/classifications/icd10/browse/2016/en. Accessed 9 Dec 2015.

[22] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laaka, B. van Ginneken, and C. I. Sánchez, "A Survey on Deep Learning in Medical Image Analysis," 2017, arXiv:1702.05747.

[23] G. Chartrand, P. M. Cheng, E. Vorontsov, M. Drozdzal, S. Turcotte, C. J. Pal, S. Kadoury, and A. Tang, "Diagnostic accuracy of digital screening mammography with and without computer-aided detection," in Radiographics, 2017, vol. 37, pp. 2113–2131.

[24] J. J. Fenton, S. H. Taplin, P. A. Carney, M. Linn Abraham, E. A. Sickles, C. D'Orsi, E. A. Berns, G. Cutter, R. E. Hendrick, W. E. Barlow, and J. G. Elmore, "Influence of computer-aided detection on performance of screening mammography." New England Journal of Medicine, 2007, vol. 356, pp. 1399–1409.

[25] C. D. Lehman, R. D. Wellman, D. S. Buist, K. Kerlikowske, A. N. Tosteson, and D. L. Miglioretti, "Diagnostic accuracy of digital screening mammography with and without computer-aided detection." JAMA Internal Medicine, 2015, vol. 175, pp. 1828–1837.

[26] B. Ghojogh and M. Crowley, "The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial," 2019, arXiv:1905.12787.

[27] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016, http://www.deeplearningbook.org.

[28] J. B. Diederik P. Kingma, "Adam: A Method for Stochastic Optimization," 2017, arXiv:1905.12787.

[29] X. Glorot and Y. Bengio, "Survey on deep learning for radiotherapy." Aistats, 2010, vol. 9, p. 249–256.

[30] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," 2016, arXiv:1602.03409.

[31] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," 2015, arXiv:1409.0575.

[32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," in Journal of Machine Learning Research, 2014.

[33] G. Ghiasi, T.-Y. Lin, and Q. V. Lei, "DropBlock: A regularization method for convolutional networks," 2018, arXiv:1810.12890.

[34] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series." in the handbook of brain theory and neural networks. MIT Press Cambridge, 1998, pp. 255–258.

[35] D. Scherer, A. Muller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in Artificial Neural Networks – ICANN. Springer Berlin Heidelberg, 2010, p. 92–101.

[36] C. Szegedy et al., "Going deeper with convolutions," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1–9.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[38] N. Tajbakhsh et al., “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,” IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1299–1312, May 2016.

[39] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” arXiv:1411.1792 [cs], Nov. 2014.

[40] “Very Deep Convolutional Networks for Large-Scale Image Recognition - Open Access Library.” [Online]. Available: http://www.oalib.com/paper/4068791#.XMM2v-gzbIU. [Accessed: 27-Apr-2019].

[41] Y. Lecun et al., “Comparison of learning algorithms for handwritten digit recognition,” International Conference on Artificial Neural Networks, Paris, pp. 53–60, 1995.

[42] A. S. Lundervold and A. Lundervold, “An overview of deep learning in medical imaging focusing on MRI,” 2018, arXiv:1811.10052.

[43] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, “Deep learning for health informatics,” in IEEE Journal of Biomedical and Health Informatics. IEEE, 2016, vol. 21, pp. 4–21.

[44] M. I. Razzak, S. Naz, and A. Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and Future,” 2017, arXiv:1704.06825.

[45] M. A. Mazurowski, M. Buda, A. Saha, and M. R. Bashir, “Deep learning in radiology: an overview of the concepts and a survey of the state of the art,” 2018, arXiv:1802.08717.

[46] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, “Dermatologist-level classification of skin cancer with deep neural networks,” in Nature, 2017, vol. 542, pp. 115–118.

[47] A. Rezvantalab, H. Safigholi, and S. Karimijeshni, “Dermatologist Level Dermoscopy Skin Cancer Classification Using Different Deep Learning Convolutional Neural Networks Algorithms,” 2018, arXiv:1810.10348.

[48] S. Hussain, S. M. Anwar, and M. Majid, “Segmentation of Glioma Tumors in Brain Using Deep Convolutional Neural Network,” 2017, arXiv:1903.05769.

[49] S. Hussein, P. Kandel, C. W. Bolan, M. B. Wallace, and U. Bagci, “Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches,” 2019, arXiv:1801.03230.

[50] P. Mlynarski, H. Delingette, A. Criminisi, and N. Ayache, “3D Convolutional Neural Networks for Tumor Segmentation using Long-range 2D Context,” 2018, arXiv:1807.08599.

[51] A. Sharma, D. Raju, and S. Ranjan, “Detection of pneumonia clouds in chest x-ray using image processing approach,” 2017 Nirma University International Conference on Engineering (NUiCONE). IEEE, 2018.

[52] S. Rajaraman, S. Candemir, I. Kim, G. Thoma, and S. Antani, “Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs,” in Applied Sciences, 2018, vol. 8, ch. 1715.

[53] S. Khobragade, A. Tiwari, C. Patil, and V. Narke, “Automatic detection of major lung diseases using chest radiographs and classification by feed-forward artificial neural network,” 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES). IEEE, 2016.

[54] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “ChestXray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases,” 2017, arXiv:1705.02315.

[55] 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.

[56] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” 2017, arXiv:1711.05225.

[57] Pan, SinnoJialin, &Qiang Yang. (2010)."A survey on transfer learning", IEEE Transactions on knowledge and data engineering, IEEE, Vol. 22.10, pp. 1345-1359.

[58] Guan, Q.; Huang, Y.; Zhong, Z.; Zheng, Z.; Zheng, L.; Yang, Y. Diagnose like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification. arXiv. 2018. Available online: https://arxiv.org/abs/1801.09927v1 (accessed on 17 June 2018).

[59]. Candemir, S.; Jaeger, S.; Palaniappan, K.; Musco, J.P.; Singh, R.K.; Xue, Z.; Karargyris, A.; Antani, S.; Thoma, G.; McDonald, C.J. Lung Segmentation in Chest Radiographs Using Anatomical Atlases with Nonrigid Registration. IEEE Trans. Med. Imaging 2014, 33, 577–590. [CrossRef] [PubMed]

[60]. Liu, C.; Yuen, J.; Torralba, A. SIFT Flow: Dense Correspondence across Scenes and Its Applications. IEEE Trans. Pattern Anal. Mach. Intell. 2011, 33, 978–994. [CrossRef] [PubMed]

[61] “NVIDIA DIGITS | NVIDIA Developer.” [Online]. Available: https://developer.nvidia.com/digits. [Accessed: 25-mar-2020].

[62] Ayan, Enes, and Halil Murat Ünver. "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning." *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*. IEEE, 2019.

[63] Asnaoui, Khalid El, Youness Chawki, and Ali Idri. "Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning." *arXiv preprint arXiv:2003.14363* (2020).

[64] Blaku, Flokart, and Nurbek Saidnassim. "Application of Deep Learning on CXR Data to Signal Cases of Pneumonia." (2018).

[65] Liu, Mao, Yumeng Tan, and Lina Chen. "Pneumonia detection based on deep neural network Retinanet." *2019 International Conference on Image and Video Processing, and Artificial Intelligence*. Vol. 11321. International Society for Optics and Photonics, 2019

Acknowledgments

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible.

In my journey towards this degree, I have found a teacher, a friend, an inspiration, a role model and a pillar of support in my Guide, Professor BO ZHANG. He has been there providing his heartfelt support and guidance at all times and has given me invaluable guidance, inspiration, and suggestions in my quest for knowledge. He has given me all the freedom to pursue my research, while silently and non-obtrusively ensuring that I stay on course and do not deviate from the core of my research and also for advanced C++ courses which he has taught me. I can safely say that I haven't learned any other engineering subject in such depth than the ones which he has taught. Without his able guidance, this thesis would not have been possible and I shall eternally be grateful to him for his assistance.

I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout every department of my life.

I would also like to thank Nankai University, and the entire staff at Nankai University, who have been so helpful and cooperative in giving their support at all times to help me achieve my goal. It would like to thank all my friends and all of my classmates, who have, in their own ways, kept me going on my path to success, assisting me as per their abilities, in whatever manner possible and for ensuring that good times keep flowing.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.