Study on COVID-19 vs Viral PNEUMONIA Detection Using Convolutional Neural Network

By
MD.Shohel Quddus (2018-2-50-028)
Moin Ahmed Rafi (2018-2-50-012)
Humayra Akter Saba (2017-1-58-005)

Supervised by
DR.Anup Kumar Paul

Department of Electronics and Communications Engineering
East West University
A/2, Jahurul Islam Avenue, Jahurul Islam City, Aftabnagar,
Dhaka-1212, Bangladesh.

September, 2022

Study on COVID-19 vs Viral PNEUMONIA Detection Using Convolutional Neural Network

Humayra Akter Saba
Student ID: 2017-1-58-005
E-mail: 2017-1-58-005@std.ewubd.edu

Moin Ahmed Rafi
Student ID: 2018-2-50-012
E-mail: 2018-2-50-012@std.ewubd.edu

MD.Shohel Quddus
Student ID: 2018-2-50-028
E-mail: 2018-2-50-028@std.ewubd.edu

Department of Electronics and Communications Engineering
East West University
A/2, Jahurul Islam Avenue, Jahurul Islam City, Aftabnagar,
Dhaka-1212, Bangladesh.

September, 2022

September, 2022
Department of Electronics and Communications Engineering
East West University

The thesis paper titled “COVID-19 Based Pneumonia vs Viral Pneumonia Detection Using Convolutional Neural Network” submitted by MD. Shohel Quddus (Student ID: 2018-2-50-028), Moin Ahmed Rafi (Student ID: 2018-2-50-012) and Humayra Akter Saba (Student ID: 2017-1-58-005) to the Department of Electronics and Communications Engineering, East West University, Dhaka, Bangladesh has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Information and Communications Engineering and approved as to its style and contents..

____

Thesis Supervisor

Dr.Anup Kumar Paul

Associate Professor

Department of Electronics and Communications Engineering

East West University, Dhaka-1212 , Bangladesh

____

Head of the Department

Mohammad Arifuzzaman Ph.D.

Chairperson , Associate Professor

Department of Electronics and Communications Engineering

East West University, Dhaka-1212 , Bangladesh

September, 2022
Department of Electronics and Communications Engineering
East West University

We undersigned hereby declare that this thesis is a presentation of our original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgment of collaborative research and discussions. The work was done under the guidance of Dr.Anup Kumar Paul, at Department of Electronics and Communications Engineering,East West University.

Humayra Akter Saba

Moin Ahmed Rafi

MD. Shohel Quddus

Acknowledgment

First of all, we would like to thank Almighty Allah for giving us enough mental and physical strength to complete our thesis properly. We would like to express our sincere gratitude to our supervisor Dr.Anup Kumar Paul for his cooperation, suggestion, guidance, and continuous encouragement through the course of the study. We are highly grateful to our Supervisor Dr.Anup Kumar Paul for reviewing this report and giving us suggestions to improve it.

Besides our supervisor, we would like to acknowledge our honorable teachers of Electronics and Communications Engineering Department for their encouraging support and discussion. We also thank our parents for their encouragement, support, and attention. We are also thankful to our classmates for their moral support which helped us to accomplish our thesis.

Dedication

To our parents, family and our honorable teachers. Both our parents and teachers gives us enough inspiration and encouragement to complete our thesis work.

Abstract

In fewer than two years, COVID-19, which is widely regarded as the most lethal virus of the twenty-first century, has been responsible for the deaths of millions of people all over the world. The novel coronavirus known as SARS-CoV-2 is the causative agent of the respiratory sickness known as COVID-19. It was first identified in Wuhan, China, in late December of 2019. According to Hopkins's projections, the virus will have killed over one million people by October 2020 and infected about 40,000,000 individuals by then. This infection has rapidly expanded across China and into other nations since then, creating a global pandemic in 2020 due to its ease of transmission from person to person via respiratory droplets. Another contagious lung disease, pneumonia is typically brought on by a bacterial infection of the alveoli. Pus occurs when infected lung tissue becomes irritated. Patients typically feel the effects of the virus in their lungs first, thus chest X-rays can help doctors diagnose the disease. Experts perform physical exams and use diagnostic tools including chest X-rays, ultrasounds, and lung biopsies to identify whether or not a patient has these illnesses. In this analysis, we recommend using a chest X-ray to prioritize people for subsequent RT-PCR testing. It would also aid in the identification of patients with a high chance of COVID and a false-negative RT-PCR who require additional testing. It is urgent to create automated technologies that could diagnose this disease in its early stages, in a non-invasive manner, and in a shorter amount of time. However, selecting the most accurate models to characterize COVID-19 patients is challenging due to the inability to compare the outputs of diverse data types and gathering methods. This is the only way to remedy the issue. As a result, much research has been conducted to establish an appropriate method for diagnosing and classifying people as COVID-19-positive, healthy, or affected by other pulmonary lung illnesses. In a few earlier scholarly works, semiautomatic machine learning techniques with limited precision were proposed.

In this study, we wanted to develop reliable deep learning approaches, which are a subset of machine learning and AI that model the way humans acquire knowledge. Data science encompasses fields like statistics and predictive modeling, two of which benefit greatly from deep learning. One component of this is what are known as convolutional neural networks (CNN). Any automatic, reliable, and accurate screening strategy for COVID-19 infection would be helpful for rapid diagnosis and reducing exposure to the virus for medical or healthcare personnel. The work takes advantage of a versatile and successful deep learning approach by employing the CNN model to predict and identify a patient as being unaffected or impacted by the disease using an image from a chest X-ray. In order to prove how well the CNN model was trained, the researchers employed a dataset consisting of 6,000 images with a resolution of 224x224 and 32 batches. Convolutional

neural networks (CNNs) were demonstrated to be very effective for medical picture classification. The authors of this piece propose using convolutional neural networks (CNNs) to automatically classify chest X-ray images for signs of COVID-19. Using the dataset, eleven current CNN models—VGG16, VGG19, DenseNet, max poling operation, and SoftMax activation function—that can distinguish between COVID-19, pneumonia, and other lung diseases—were first used to identify the symptoms of COVID-19. To avoid overfitting, we used a stratified 5-fold cross-validation approach, allocating 90 percent of the dataset to training and 10percent to testing (unseen folds), and validating our model on 20 percent of the training data. A 95 percent accuracy rate was achieved during performance training with the trained model. Python's built-in machine learning functionality utilizes a confusion matrix. The predictions made by a classification issue are recorded in a confusion matrix. For each category, the number of correct and incorrect predictions is represented by a count value. That's the key to deciphering the matrix of ambiguity. The confusion matrix illustrates how your classification model generates predictions despite the uncertainty it faces. The research study can use chest X-ray pictures to identify and detect COVID-19, normal, and pneumonia infections, according to the results of the tests.

Table of Contents

List of Figures

List of Tables

Chapter I

Background

1.1 Introduction

The novel coronavirus (COVID-19) is a previously unidentified and undiagnosed virus in humans. On March 11, 2020, the World Health Organization (WHO) declared COVID-19 pneumonia a pandemic; it is a contagious respiratory disease that is altering the world. The majority of patients infected with COVID-19 suffered from moderate to severe illnesses, such as asthma, and a number of them also suffered from pneumonia which may have been fatal. Despite the fact that the prevalence of infection in children and young adults is lower, there are assumptions that elderly individuals with basic clinical issues such as cardiovascular disease, diabetes, persistent respiratory infection, renal or hepatic diseases, and malignant growth are doomed to develop the true disease (Dong et al., 2020). Both the number of transmission cases and the mortality rate are on the rise, resulting in a significant increase in intensive care unit (ICU) admissions and a pressing need for rapid and accurate screening methods. According to sources, one of the most crucial phases in the fight against COVID-19 is the early discovery of infected individuals, followed by the adoption of treatment protocols in severe instances and quarantine measures to prevent disease transmission (Polat et al., 2021). Over 18 million COVID-19 infections have been detected in 213 countries across the globe.[1]

Lack of a correct diagnosis can lead to fatal pneumonia, which is a potentially fatal condition. It's a severe form of respiratory illness brought on by infectious agents like viruses or bacteria. Breathing it in can cause damage to your lungs, and it can spread through your nose or throat. Droplets released into the air during a sneeze or cough might potentially spread the virus. Pneumonia causes decreased oxygen intake and uncomfortable, labored breathing due to inflammation of the bronchial tubes produced by fluid buildup or mucopurulent secretions. Over a million individuals worldwide have lost their lives to pneumonia, primarily those over the age of 50 and those under the age of five who have weaker immune systems. Nearly 58,000 people lost their lives to it in the Philippines in 2016, making it the third greatest cause of death after cardiovascular disease and cancer. In a society where hospitals are overcrowded, a misdiagnosis of pneumonia is a real possibility due to the similarity of 19 symptoms and indicators. The costly setbacks in treatment and the possibility of coming into contact with other COVID-19-positive patients that can result from misdiagnosing pneumonia or other non-COVID-19 diseases as COVID-19 infections are real concerns for COVID-19-positive patients. [2]

Detecting and isolating patients infected with COVID-19 as soon as their identity is established is the proposed method for reducing the epidemic transmission rate and flattening the transmission curve. However, professionals are currently conducting continuing research and clinical trials for potential therapies. The World Health Organization has designated Reverse Transcription Polymerase Chain Reaction (RT-PCR) as the laboratory gold standard for COVID-19 diagnosis. This is the most popular approach for identifying COVID-19, however, it is a laborious, difficult, arduous, and time-consuming process with a positive rate between (60 percent) and (70percent). The image-based COVID-19 diagnostic has proven to be effective in terms of rapid and accurate identification. When it comes to diagnostic and detection tactics, however, a significant proportion of recent material has focused on radiology techniques such as portable chest radiography (X-rays). 2020 (Chung et al.; Zhou et al.). Every day, several devices are produced and improved in order to enhance image processing. It is a medical image directory structure divided into three subfolders containing Chest X-ray (CXR) images (COVID, NORMAL, and PNEUMONIA).

COVID: 576 images
NORMAL: 1583 images
PNEUMONIA: 4303 images

Each picture goes through preprocessing to decrease it to 224x224 pixels in PNG format. Several countries are now developing COVID-19 vaccines. The WHO has authorized the use of a number of vaccines, including those manufactured by Pfizer, AstraZeneca, Moderna, Serum Institute of India Pvt. Ltd., Janssen, Sinopharm, and Sinovac. The disease's prevalence has been drastically reduced as a result of these permitted vaccines. [3]

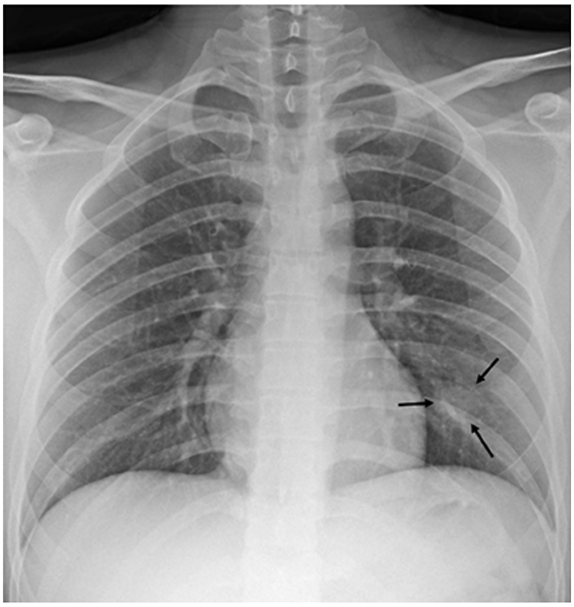

Figure 1.1: COVID-19 X-ray.

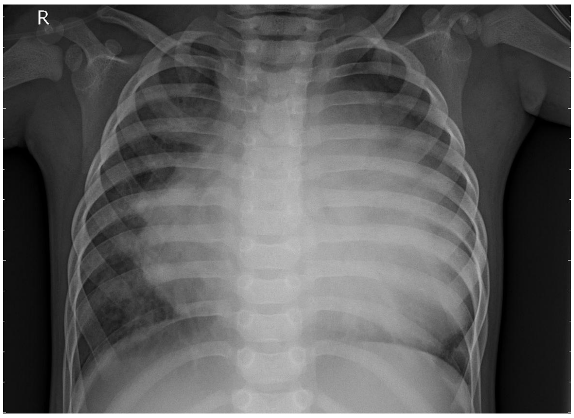

Figure 1.2: PNEUMONIA X-ray.

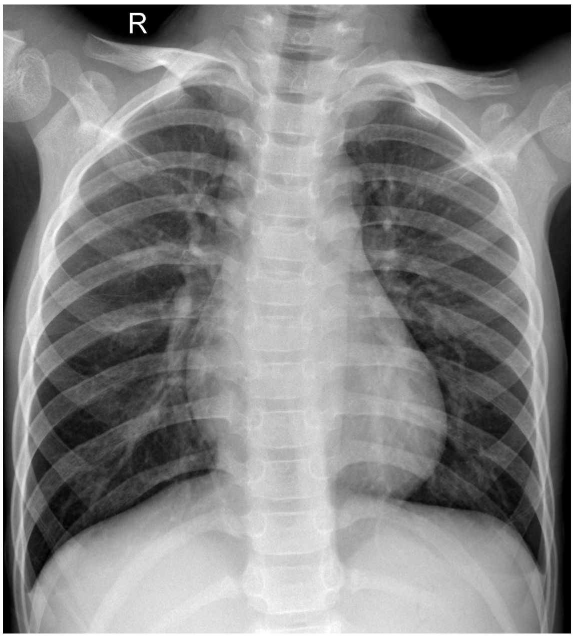

Figure 1.3: Normal X-ray.

1.2 Artificial Intelligence (AI)

Deep learning-based Artificial Intelligence (AI) solutions are presently employed to solve a variety of biological concerns, such as the detection of breast cancer and brain tumors. It is a member of the machine learning techniques family. Currently, artificial intelligence is being used to automate the diagnosis of a number of diseases, having demonstrated its efficacy and strong performance in automated image categorization challenges using a variety of machine learning approaches. DL is a branch of AI and ML that increases the performance of AI and ML applications. Deep neural networks (DNNs) have been proposed by researchers to aid in the detection of the disease on Chest X-ray images (Wang and Wong (2020), Shoeibi et al. (2020)). For the COVID-19 images, Wang and Wong (2020) reported a test accuracy of 92.6 percent, a recall of 96.4 percent, and a precision of 87Percent. [4].

1.3 Convolutional Neural Networks (CNN):

The CNN model is used in a number of influential publications to diagnose COVID-19 pneumonia. CNN links. The four essential layers are the convolutional layer, the rectified linear unit (ReLU) activation layer, the subsampling layer, and the fully-connected layer. In addition, CNNs may be trained on datasets of any size by repeatedly iterating over large or small data batches. Other recent improvements include the novel convolutional block attention module (Zhang et al., 2021), data augmentation techniques, new architectures and pre-trained models (He et al., 2016; Nayak et al., 2021), optimization algorithms (Kingmaand Ba, 2014), activation functions (Pedamonti, 2018), regularization techniques such as dropout (Srivastava et al., 2014), and batch normalization. CNNs can be included into limited-capacity devices when their size is appropriate, allowing untrained users to use these devices to identify a range of diseases in undeveloped regions (Rong et al., 2020). The suggested model was trained, validated, and evaluated using retrospective CXIs from real-world patients with COVID-19 pneumonia and other forms of pneumonia. The trained CNN is able to evaluate new photographs by recognizing patterns in each image that indicate to a specific disease. Transfer learning was used to train the neural network to distinguish between normal lungs, COVID-19, and pneumonia. Open datasets from COVID-19, pneumonia, and normal chest X-rays were combined to generate our COVID-19 dataset. It is feasible that the development of a reliable automated COVID-19 detection technology will allow clinicians to receive a second opinion and lessen their workload.

1.4 Purpose

Using chest X-ray images and CNNs, we will introduce novel models for detecting COVID-19 and other types of pneumonia. The two classification situations covered by the proposed models are binary classification (COVID-19 vs. Non-COVID) and three-class classification, and they were designed to provide accurate diagnoses in more output classes than previous studies (COVID-19 vs. Normal vs. Pneumonia). To improve accuracy and prevent overfitting, two steps should be taken: (1) enlarging the data set (considering additional data repositories and employing data augmentation techniques) while balancing classification scenarios; and (2) incorporating regularization techniques such as dropout and automating hyperparameter optimization.

1.5 Aim of proposal

A reader suggested a hybrid artificial intelligence solution to improve the accuracy of models used to identify COVID-19 patients from chest X-ray pictures. This system utilized machine learning and deep learning algorithms, notably Convolutional Neural Network (CNN) with softmax classifier. COVID-19 is frequently detected in association with pneumonia symptoms, which genetic and imaging studies might reveal. Rapid detection of COVID-19 can aid in disease containment. By rapidly recognizing COVID-19, image tests can aid in its containment. Therefore, they concluded that the proposed model had an accuracy rate of 95.7 percent. The purpose of this study is to improve the diagnostic accuracy of Covid-19 and pneumonia's using CNN. We have utilized a variety of evaluation methods to demonstrate the superior performance of the suggested model.

1. In order to halt the spread of COVID-19, it is important to review a substantial number of reported cases for proper isolation and treatment.

2. Using CNNs, we can differentiate between pictures of pneumonia and normal lung tissue in lung scans that are positive or negative for the Covid-19 protein.

3. The outcomes of our trials demonstrate the viability of models for COVID-19 detection, prevention, and control. The overall accuracy of the suggested model is 95.75 percent on average. rachmi2016stunting

Chapter II

Literature Review

2.1 Background

There have been a number of research done to emphasize the scientific facts and design considerations of COVID-19. Several techniques for recognizing, analyzing, interpreting, and settling conclusions from visual data are shown here. In a variety of fields, the notion of deep learning (DL) has been put forth and has shown promising accuracy results. In order to facilitate the analysis and development of AI systems for COVID-19 identification, Cohen et al. recently compiled a repository of CT and chest x-ray images from cases of Middle East Respiratory Syndrome (MERS), Acute Respiratory Distress Syndrome(ARDS), and SARS. The immediate release of the proposed COVID-19 data led to the development of several automated diagnostic algorithms based on chest x-ray images. [5] Consequently, it is essential to underline that the COVID-19 dataset continues to grow as new patient cases are regularly added and frequently made public. Wang et al. presented a DL architecture known as COVID-19 in which writers utilize the COVID-19 public dataset and the open dataset of Chest x-ray images (Pneumonia). COVIDx is a dataset of 6,000 posteroanterior chest radiographs from 576 patient instances that was derived by the author. Their investigation focuses on COVID-19, Normal, and Pneumonia image categories.

2.2 Histology Image Data in Deep learning

Histology is the science of the microscopic nature of organisms' cells and tissues. The word "histology" was first used by German anatomist and physiologist Karl Mayer in 1819 in his work "On histology and a new category of tissues of the human body." Pathologists analyze tissue under various magnifications of a microscope slide to spot morphological traits that can be used to diagnose diseases like cancer and Covid-19. Histopathology images can be captured using specialized cameras and a microscope, followed by an automated computerized procedure. The biopsy specimen is embedded in wax and dyed with one or more stains to study the architecture and constituents of tissues under a microscope. Computational pathology goes beyond the simple recognition of morphological patterns, even though the majority of AI research is still concentrated on the detection and grading of tumors and various types of disorders in digital histology and radiology. Implementing AI-based computer-aided medicine combined with clinical data from EHR, including people's clinical risk factors of human-to-human contacts and a variety of different social data, may give quick control of this public health issue with

a greater level of quality and safety. Histology images aid in identifying the various types of cell nuclei and their architecture based on a pattern. Histopathologists examine the regularities of cellular architecture and tissue distributions to assess cancerous and other types of disease regions such as-Covid-19infected lungs, as well as the degree of malignancy. A number of AI businesses are developing products to combat the COVID-19 pandemic. For instance, the polymerase chain reaction (rt chain reaction (RT-PCR) results, imaging tests, and Capital by investing Inspection, Korea's (http://www.jlk-inspection.com//medical/main) universal AI system, AIHuB, are all being integrated to offer COVID-19 diagnosis. In order to identify and notify patients who are likely COVID-19 positive, Persivia, Massachusetts (https://persivia.com/covid-19-detection/), has introduced a new monitoring module based on its Soliton AI engine. An analytical platform called Biovitals Sentinel was created by Biofourmis, a Massachusetts-based company (https://www.biofourmis.com/). It uses histological images based on artificial intelligence to detect early patient deterioration and enable earlier treatments. Machine learning could considerably improve the efficiency and effectiveness of randomized clinical trials for COVID-19, according to Schaar et al., It has the potential to speed up subject recruitment from distinguishable subgroups and subject assignment to treatment or control groups, as well as drastically lower error and need a great deal fewer individuals. [6]
Even though machine learning has achieved encouraging results and offers various advantages in computational pathology, the following difficulties must be tackled before deep machine learning may be applied in the clinical setting. Histology is not that difficult to understand; lab experience is required. It requires patience but is not hard.

2.3 Image processing of Convolutional Neural Network(CNN)

Deep learning algorithms called convolutional neural networks are extremely effective at analyzing images. You will learn how to create, train, and assess convolutional neural networks in this article. This edition includes a new unique network made up of convolution and pooling layers in place of entirely connected hidden layers. CNN operates by detecting characteristics from images. A CNN is made up of the following components:

1. A monochrome image serves as the input layer.
2. The binary or multi-class labels that make up the output layer.
3. Convolution, ReLU (rectified linear unit), pooling, and a fully connected neural network are the components of the hidden layers. Right now, these are the fastest algorithms available for automatically processing images. These algorithms are widely used by businesses to do tasks like item identification in images. [7]

2.4 Three Layers of CNN

Image and video recognition applications use convolutional neural networks. The primary applications of CNN are in image analysis tasks including segmentation, object identification, and picture recognition. Convolutional Neural Networks include three different sorts of layers:

1. Convolution Layer.
2. Pooling Layer.
3. Fully Connected Layer.

2.5 Limitations and Motivation

A timely diagnosis is essential due to the rapid spread of the coronavirus and the severe effects it has on humans. As mentioned, pneumonia affects a large number of people, especially children, in developing and undeveloped nations with overcrowding, inadequate sanitation, hunger, and a lack of medical care. Early diagnosis is key to curing pneumonia. Autonomous CAD with generalization capability is needed to detect the condition. Most prior strategies in the literature focused on building a single CNN model for pneumonia case categorization, and ensemble learning has not been addressed. COVID-19 results are often inconspicuous. Radiologists can identify 65 percent of positive cases. AI tools may assist mitigate this disadvantage. Doctors will have an X-ray-based early warning tool for COVID-19. CNN's achievements in identifying disease are encouraging, but X-ray-trained models from one hospital or set of hospitals have not been shown to work in other institutions. Current constraints include photo collection biases. If CNN performance predictions are based on CXR test data, they may exaggerate clinical performance. The location of acquisition may be anticipated with great accuracy, both in terms of the CXR equipment used and the hospital department. When creating these models, it's vital to consider that the network may learn the source of the images rather than the disease. Generalization is inversely proportional to the amount of data (images) used to train the algorithm. However, this isn't always the case due to possible biases coming from an imbalance in the numbers of positive and negative images used for training, which are typically of different origin, and the varied features of the images in each set, such as varying toma, pulse width, detection shape, image size, pixel intensity, artifacts, and labels, among others, which, if not handled appropriately, can provide inaccurate findings. Due to the limits of the PCR approach, CNN architectures were examined to overcome this issue. A model for Covid-19 identification was presented that automatically learns characteristics from photos. The proposed model was tested on two publicly available chest X-ray datasets and the Pneumonia Detection Challenge dataset using three-fold cross-validation. The findings outperform state-of-the-art procedures, making the method real-world feasible.As mentioned, pneumonia affects a large number of people, especially children, in developing and undeveloped nations with overcrowding, inadequate sanitation, hunger, and a lack of medical care. Early diagnosis is key to curing pneumonia. Autonomous CAD with generalization capability is needed to detect the condition. Most prior strategies in the literature focused on building a single CNN model for pneumonia case categorization, and ensemble learning has not been addressed. COVID-19 results are often inconspicuous. Radiologists can identify 65 percent of positive cases. AI tools may assist mitigate this disadvantage. Doctors will have an X-ray-based early warning tool for COVID-19. CNN's achievements in identifying disease are encouraging, but X-ray-trained models from one hospital or set of hospitals have not been shown to work in other institutions. Current constraints include photo collection biases. If CNN performance predictions are based on CXR test data, they may exaggerate clinical performance. The location of acquisition may be anticipated with great accuracy, both in terms of the CXR equipment used and the hospital department. When creating these models, it's vital to consider that the network may learn the source of the images rather than the disease. Generalization is inversely proportional to the amount of data (images) used to train the algorithm. However, this isn't always the case due to possible biases coming from an imbalance in the numbers of positive and negative images used for training, which are typically of different origin, and the varied features of the images in each set, such as varying toma, pulse width, detection shape, image size, pixel intensity, artifacts, and labels, among others, which, if not handled appropriately, can provide inaccurate findings. Due to the limits of the PCR approach, CNN architectures were examined to overcome this issue. A model for Covid-19 identification was presented that automatically learns characteristics from images. The proposed model was tested on two publicly available chest X-ray datasets and the Pneumonia Detection Challenge dataset using three-fold cross-validation. The findings outperform state-of-the-art procedures, making the method real-world feasible. [5]

2.6 Literature Review

According to the literature cited above, X-Ray pictures have been used for the majority of the work. In current history, researchers have started examining and analyzing chest X-ray images to discover COVID-19 using deep learning approaches. This research does a thorough analysis of the deep learning techniques that are currently in use for coronavirus infection identification using CXR pictures. Despite the fact that there are additional surveys in the literature, most of them have a larger focus. include AI in computational biology and medicine, data science techniques for pandemic modeling, AI with the Internet of Things (IoT), AI for text mining and natural language processing (NLP), and AI in medical image processing. Ulhaq et al., for instance, reviewed every technique used to combat coronaviruses. This gives a broad overview of what is going on in the world of research. The segmentation of lung images was discussed in a survey on the use of computer vision techniques for COVID-19. During the process of this research, we might well be able to detect the early stages of lung and heart disease. This paper's only focus is on deep learning-based coronavirus detection techniques. This publication reviews all approaches mentioned in the literature in the hopes of assisting researchers in developing improved coronavirus detection methods. Three databases may be used to get chest X-ray images: Chest X-ray Images (COVID-19), Chest X-ray Images (Normal), and Chest X-Ray Images (Pneumonia). The model is trained on a small dataset that includes normal (1266), pneumonia (3418), and Covid-19 (460) pictures, as well as test data that includes normal (317), pneumonia (885), and Covid-19 (116) images. Their suggested model's total accuracy is 95.75 percent. The authors also employed deep learning to identify COVID-19 patients from a small number of chest X-ray images, which worked effectively. Employing a few, selected datasets for the study of cell biology to test algorithmic problems. In this work, typical metrics for assessment and comparison are presented together with the techniques, datasets, and utilized datasets. Future directions are also detailed. [3]

Chapter III

Methodology

3.1 Convolutional Neural Network Architecture

The ConvNet's task is to compress the images into a more manageable format while maintaining components that are crucial for getting a good prediction. This is crucial for creating an architecture that can learn features and scale to enormous datasets. Three

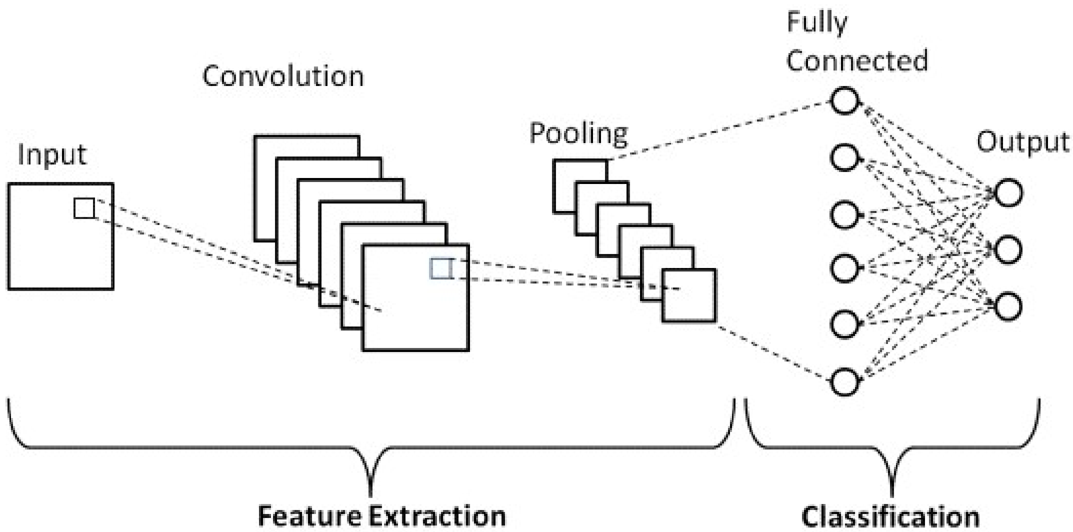

Figure 3.1: A convolutional neural network

layers make up a convolutional neural network, or ConvNets in short. Let's take a closer look:

3.2 What Is a Convolution?

Convolution is an method where two sources of information are knotted; it's an operation that changes a function into something else. Convolutions have been used for a long time usually in image processing to blur and sharpen images, but it can also to perform for other operations. (e.g. enhance edges and emboss) CNNs enforce a local connectivity pattern between neurons of adjacent layers. CNNs make use of filters (also known as kernels), to detect the features throughout an image.

1. Convolution

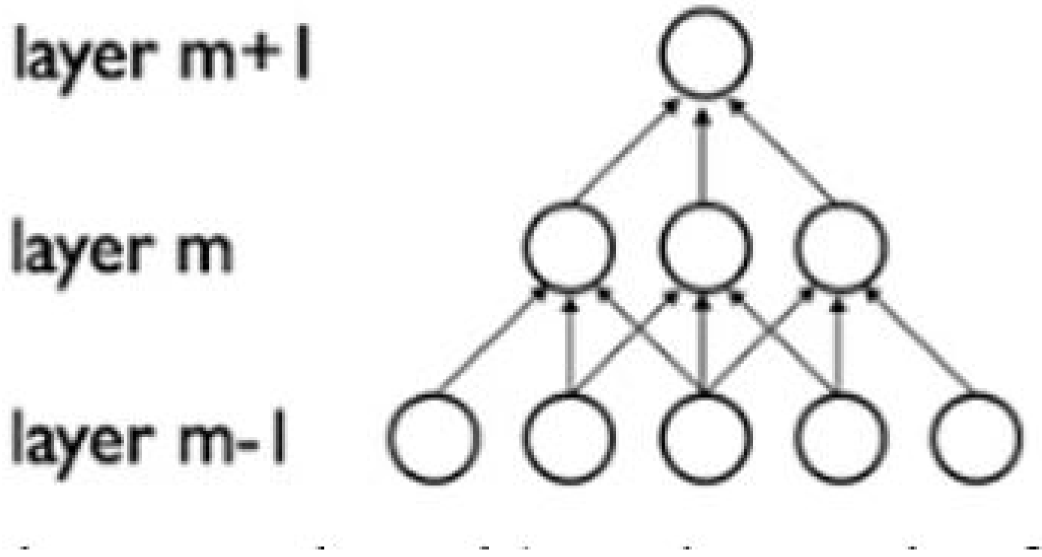

Figure 3.2: A demo CNN layer.

2. Non Linearity (ReLU)
3. Pooling or Sub Sampling
4. Classification (Fully Connected Layer) [8]

A convolutional neural network's first layer is always a convolutional layer. Convolutional layers operate on the input using a convolution operation and send the outcome to the following layer. All the pixels in a convolution's receptive area are combined into a single value. For instance, if you apply a convolution on an image, you will both reduce the size of the image and combine all of the field data into a single pixel. A vector is the convolutional layer's final output. We can employ several types of convolutions depending on the problem type we need to solve and the features we want to learn.

3.3 Proposed Convolutional Neural Network

The three fundamental ideas of sparse interaction, parameter sharing, and equivariant representation, which motivated computer vision researchers, are all used in convolution. Let's explore each of them in more detail. In simple neural network layers, matrix multiplication is used to describe how an input unit interacts with an output unit using a matrix of parameters. This suggests that all input and output devices are in communication with one another. However, there is little interaction between convolution neural networks. Making the kernel smaller than the input does this; for example, an image may have millions or thousands of pixels, but by processing it with the kernel, we can identify important information that is only of tens or hundreds of pixels. This suggests that we should store fewer parameters, which reduces the model's memory requirement and increases the statistical power of the model. If computing a specific attribute at a specific spatial point (x1, y1) is advantageous, then it should also be advantageous elsewhere, such as (x2, y2). It suggests that while building an activation map for a single two-dimensional slice or for a single activation map, neurons must use the same set of weights. In contrast to a typical neural network, where each component of the weight matrix is only ever used once, a convolution network incorporates shared parameters, meaning that weights applied to one input are the same as weights employed elsewhere for getting output. As a result of parameter sharing, the layers of a convolution neural network will have the property of suitable balance to translation. It states that the output will change in accordance with how the input is changed..Convolutional neural networks (CNNs) are a type of artificial neural network that are used for image recognition and processing. CNNs are specifically designed to process pixel input. CNNs are powerful artificial intelligence (AI) image processing systems that use deep learning to perform both generative and descriptive tasks. They commonly employ machine vision, which incorporates image and video recognition, recommendation engines, and natural language processing (NLP). For image processing, traditional neural networks must be given images in low-resolution, pixel-by-pixel chunks. A neural network is a type of hardware or software system that is based on how neurons in the human brain work. The structure of CNN's "neurons" is more like to that of the frontal lobe, which processes visual information in humans and other animals. The layers of neurons address the problem of piecemeal picture processing in traditional neural networks by covering the entire visual field. A CNN employs a multilayer perceptron-like system that has been tuned for minimal processing requirements. A CNN has 3 layers: an input layer, an output layer, and a hidden layer, as well as several convolutional, pooling, fully connected, and normalizing layers. The removal of constraints and increase in efficiency for image processing leads to a system, which is much more efficient and simple to train for image processing and natural language processing.

3.4 Convolutional Layer (CONV)

The first layer of a convolutional neural network is always a convolutional layer. Convolutional layers perform a convolution operation on the input before sending the results to the next layer. The receptive area of a convolution combines all of the pixels into a single value. Convolution, for instance, can be used to both shrink the size of an image and integrate all of the field data into a single pixel. The output of the convolutional layer is a vector. Depending on the kind of problem we need to solve and the features we want to learn, we can use a variety of convolutional techniques.

3.5 Pooling Layer

The task of decreasing dimensionality falls on this layer. It helps to lessen how much computing power is needed to process the data. Maximum and average pooling are the two categories into which pooling can be classified. Max pooling returns the highest value from the kernel-covered region of the image. Average pooling gives the average of all the values in the area of the picture covered by the kernel.

Why Do We Need Pooling in a CNN?

The fundamental units of a convolutional neural network used for computer vision applications like image recognition are called convolutional layers. An image is passed through a filter in a convolutional layer to extract features, creating a feature map that may be fed into the following convolutional layer to extract higher-level information. CNNs can recognize increasingly complex structures and objects in a picture by stacking numerous convolutional layers.

Convolutional layers have a significant issue in that the feature map generated by the filter is location-dependent. In other words, convolutional neural networks train themselves to link the existence of a certain feature to a certain area in the input image. Performance may be significantly affected by this. Instead, we prefer translation invariant feature maps and networks (a fancy expression that means that the location of the feature should not matter). A convolutional neural network's feature map created by an earlier convolutional layer and a non-linear activation function is typically used to apply pooling.

How Does Pooling Work?

The pooling process' fundamental steps are quite similar to those of the convolution operation. You pick a filter and move it over the output feature map of the convolutional layer that came before. The most popular filter size is 22, and a stride of 2 is used to slide it across the input. The pooling filter determines an output on the receptive field based on the sort of pooling operation you've chosen (the part of the feature map under the filter). There are various methods for pooling. The two methods that are most frequently employed are max-pooling and average pooling.

3.6 Max Pooling

The filter in max pooling merely chooses the highest pixel value present in the receptive field. For instance, you would choose 9 if the field had 4 pixels with the values 3, 9, 0, and 6.

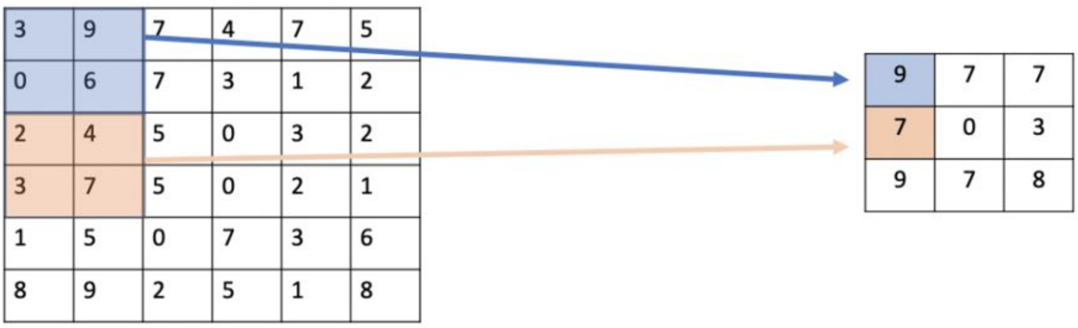

Figure 3.3: Maxpooling.

3.7 Average Pooling

The average pixel values in the receptive field are determined through average pooling. The average pooling layer would generate an output of 4.5 given 4 pixels with the values

3,9,0, and 6. Rounding to the next whole number gets us 5.

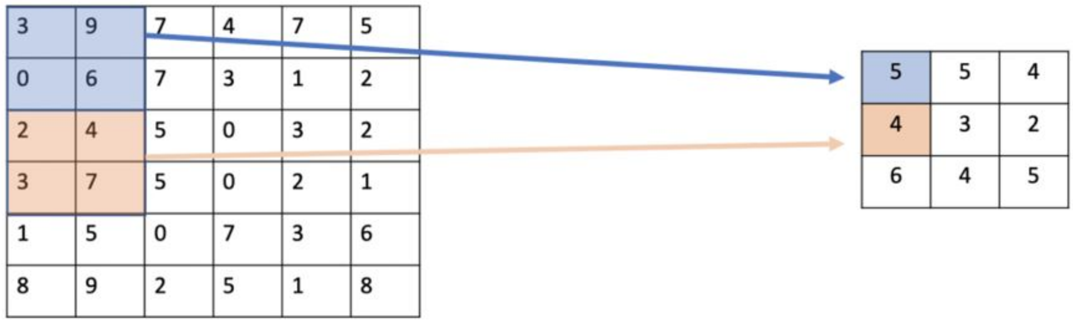

Figure 3.4: Average pooling.

3.8 Understanding the Value of Pooling

You can conceive of the numbers that the pooling layers compute and store as denoting the presence of a specific feature. The neural network's capacity to detect the feature would be based on the feature's placement on the original feature map. The network would learn to identify the characteristic associated with the number 9 with the upper left quadrant, for instance, if the number 9 was only detected in the upper left quadrant. Pooling is used to extract this feature into a more generic, smaller map that merely shows if a feature is present in that specific quadrant or not. The map gets smaller with each succeeding layer, retaining only the crucial details regarding the existence of the interesting features. The map becomes less dependent on the position of the feature as it gets smaller. The feature should seem similarly reflected on the map created by the pooling layers as long as it has been discovered in the general area of the original site. Max pooling is aware of the more pronounced features and edges in the receptive field because of its emphasis on extreme values. However, average pooling provides averages rather than picking the most extreme values, which results in a smoother feature map. Max pooling is used more frequently in practice since it is generally more effective at finding salient features. Average pooling is only used to shrink feature maps to a specific size in real-world applications. Pooling, which can compress feature maps, can also assist in categorizing photos of various sizes. In a neural network, the classification layer anticipates receiving inputs in the same format. As a result, we typically feed photos that are the same size. We can summarize photos of various sizes while still producing feature maps of corresponding sizes by changing the offsets throughout the pooling phase. Generally speaking, pooling is very useful for image classification tasks where you only need to find the presence of a specific object in an image but don't care exactly where it is. The fact that pooling filters have a longer stride than convolutional filters and produce smaller outputs contributes to the network's effectiveness and speeds up training. So location invariance can significantly boost the network's statistical effectiveness.

3.9 Fully Connected Layer (FC)

The flattened input used by the fully connected layer (FC) means that each input is tied to each neuron. The flattened vector is then transmitted via a few more FC layers, where the usual mathematical functional operations are carried out. At this moment, the classification process begins. If FC layers are present, they are often located near the end of CNN architectures.

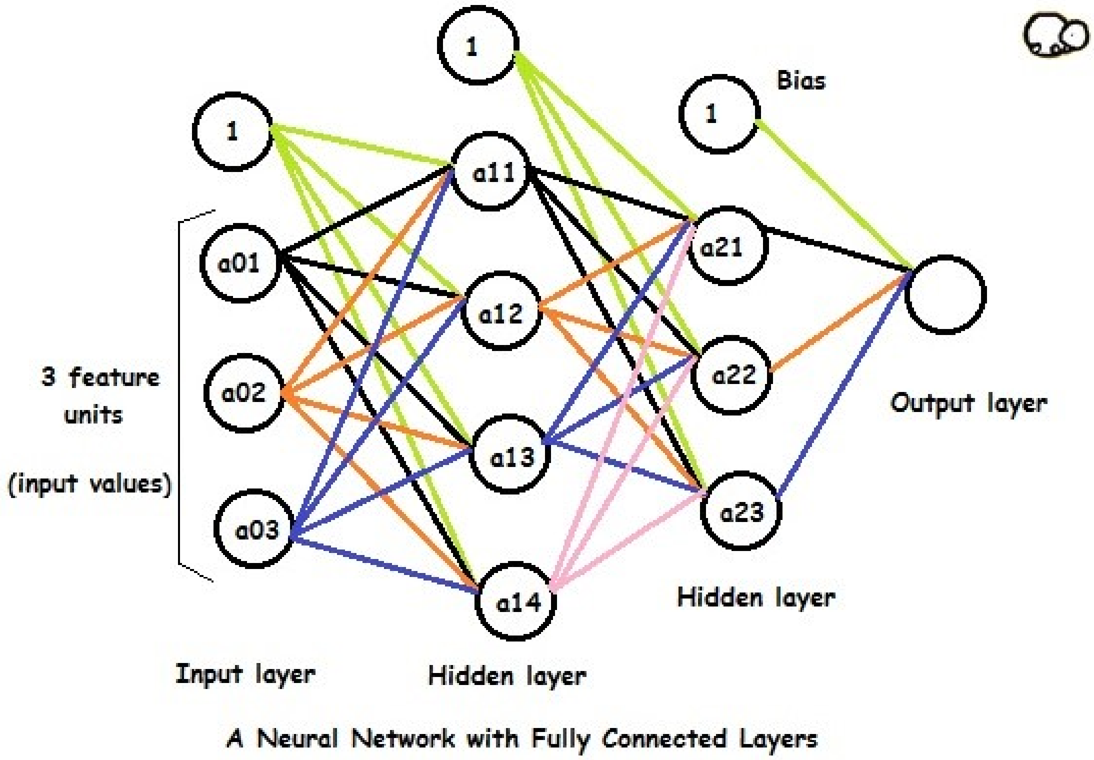

Figure 3.5: Fully Connected Layer.

What is Fully Connected Layer

1. A layer of an artificial neural networks where each element of the layer is connected to each element of the following layer. Learn more in: Enhanced Footsteps Generation Method for Walking Robots Based on Convolutional Neural Networks
2. A network layer where all neurons of the layer are connected to all neurons of the previous layer. Learn more in: Convolutional Neural Network
3. A network layer where all neurons of the layer are connected to all neurons of the previous layer. Learn more in: Deep Learning on Edge: Challenges and Trends. [9]

3.10 Activation Functions

An activation function in a neural network describes how a node or nodes in a layer of the network translate the weighted sum of the input into an output. A "transfer function" is another name for the activation function. It may be referred to as a ”squashing function” if the output range of the activation function is constrained. Numerous activation functions have nonlinear behavior, which is referred to as ”nonlinearity” in network or layer design. Different activation functions may be used in different regions of the model, and the choice of activation function has a significant impact on the neural network’s capacity and performance. Although networks are built to utilize the same activation function for all nodes in a layer, technically the activation function is applied before or after the internal processing of each node in the network. A network may have three different kinds of layers: output levels that produce predictions, hidden layers that receive data from one layer and transfer it to another, and input layers that take raw input directly from the domain. Typically, the same activation function is used by all buried levels. The sort of prediction needed by the model will determine what activation function is used in the output layer, which is often different from the hidden layers. The activation function is applied before or after each node in the network’s internal processing, despite the fact that networks are designed to use the same activation function for all nodes in a layer. Three different layers can exist in a network: input layers that receive raw data from the domain, hidden layers that accept data from one layer and pass it to another, and output levels that provide predictions. All buried levels typically employ the same activation function. The activation function employed in the output layer, which is frequently different from the hidden layers, will depend on the type of prediction required by the model.

3.11 The Dropout layer

Another significant CNN feature is a dropout layer. The Dropout layer serves as a mask, keeping all other neurons viable but removing some neurons’ contributions to the following layer. If the input vector is given a Dropout layer, However, if we apply it to a hidden layer, some hidden neurons are also destroyed, while some of its properties are lost . [10]

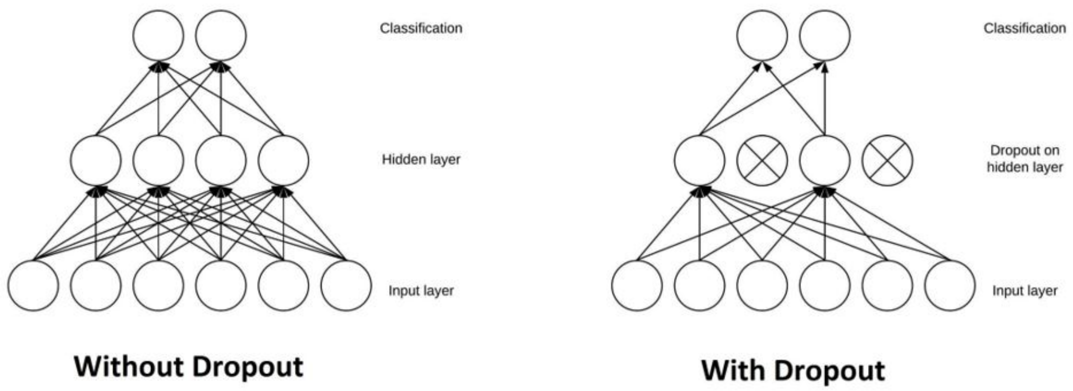

Figure 3.6: Dropout Layer.

Because they avoid overfitting on the training data, dropout layers are crucial in the training of CNNs. If they are absent, the first set of training samples has an excessively large impact on learning.

3.12 Gradient Descent

An established optimization algorithm is gradient descent. This optimization approach is typically the first one used to train machine learning. To better understand how the phrase ”Gradient Descent” applies to machine learning algorithms, let’s break it down. A gradient is a calculation that expresses how steep a line or curve is. It describes the direction of the climb or descent of a line mathematically. Descent is the activity of going downwards. Therefore, using the two straightforward meanings of these words, the gradient descent algorithm quantifies downward motion. You try to find the weights and biases in the network that will aid you in solving the problem at hand when you train a machine learning algorithm. You might, for instance, struggle with categorisation. You should be able to tell whether a picture is of a dog or a cat when you are looking at it. You train your algorithm using training data using samples of correctly labeled cat and dog photographs in order to create your model.

Although the aforementioned example was of classification, the issue could be one of localisation or detection. However, the effectiveness of a neural network in solving a problem is portrayed as a function, more precisely as a cost function, which assesses how inaccurate a model is. The weights and biases used for the final model are influenced by the partial derivatives of the cost function. The process known as gradient descent makes it easier to find parameter values that reduce the cost function to a local minimum or to the best possible accuracy. [11]

3.13 Adam

Adam is distinct from conventional stochastic gradient descent. Stochastic gradient descent maintains a consistent learning rate (referred to as alpha) that stays the same throughout training for all weight updates. A learning rate is maintained and independently changed for each network weight as learning advances (parameter). Adam is combining the benefits of two additional stochastic gradient descent modifications. Specifically: [12]

What is the Adam optimization algorithm?

In place of the conventional stochastic gradient descent method, Adam is an optimization technique that may be used to iteratively update network weights depending on training data.

How Does Adam Work?

Compared to traditional stochastic gradient descent, Adam is unique. For all weight updates, stochastic gradient descent maintains a constant learning rate (referred to as alpha), which does not change throughout training. As learning progresses, a learning rate is maintained and independently adjusted for each network weight (parameter). The authors describe Adam as combining the advantages of two other extensions of stochastic gradient descent. Specifically:

Adaptive Gradient Algorithm (AdaGrad)

Performance on issues with sparse gradients is enhanced by (AdaGrad), which keeps a per-parameter learning rate (e.g. natural language and computer vision problems).

Root Mean Square Propagation (RMSProp)

(RMSProp) that also keeps per-parameter learning rates that are adjusted depending on the mean of recent weight gradient magnitudes (e.g. how quickly it is changing). The algorithm performs effectively on online and non-stationary issues, according to this.

Adam understands the advantages of both RMSProp and AdaGrad. Adam also uses the average of the second moments of the gradients, as opposed to adjusting the parameter learning rates based on the average first moment (the mean), as in RMSProp (the uncentered variance). Adam is aware of the benefits of RMSProp and AdaGrad. Additionally, Adam adjusts the parameter learning rates based on the average second moments of the gradients rather than the average first moment (the mean), as in RMSProp (the uncentered variance). The pace at which the method calculates an exponential moving average of the gradient and the squared gradient is controlled by the parameters beta1 and beta2. [12]. Due to the initial value of the moving averages and the recommended beta1 and beta2 values of close to 1.0, moment estimations are biased towards zero. By first computing the inaccurate estimates and then the corrected estimates for bias, this bias is lessened.

3.14 Edge Detection

Essentially, edges in an image are abrupt shifts or discontinuities. Edges are defined as significant transitions in an image. The initial stage of picture recognition is edge detection. Because it delineates and separates a plane, object, or aspect from other objects, an edge is useful. A border is made up of pixels that have varied gray tone intensities from their neighboring pixels. Edges of an image provide shape information. Therefore, the initial stage in image recognition is to look for a picture's edges. Types of edges • Horizontal edges • Vertical Edges • Diagonal Edges There are two methods for detecting edges: gradient-based and Laplacian-based. The first order derivative of the image is used when creating a gradient. The broad margins and noise sensitivity of the first order derivatives make them undesirable. Images created using second order derivatives in the Laplacian approach.An edge is found, which indicates a major spatial change. Although more sophisticated approaches to automated edge identification, 2nd Order Derivative operators are still relatively noise-sensitive.

3.15 Backpropagation

The foundation of neural network training is backpropagation. It is a technique for adjusting a neural network's weights based on the error rate recorded in the previous epoch (i.e., iteration). By properly tweaking the weights, you may lower error rates and improve the model's reliability by broadening its applicability. The term ”backward propagation of errors” is shortened to ”backpropagation” in neural networks. It is a common technique for developing artificial neural networks. With regard to each weight in the network, this technique aids in calculating the gradient of a loss function.

How Backpropagation Algorithm Works?

The gradient of the loss function for a single weight is calculated by the neural network's back propagation algorithm using the chain rule. In contrast to a native direct calculation, it efficiently computes one layer at a time. Although it computes the gradient, it does not specify how the gradient should be applied. It broadens the scope of the delta rule's computation. [13] Consider the following Back propagation neural network example diagram to understand:

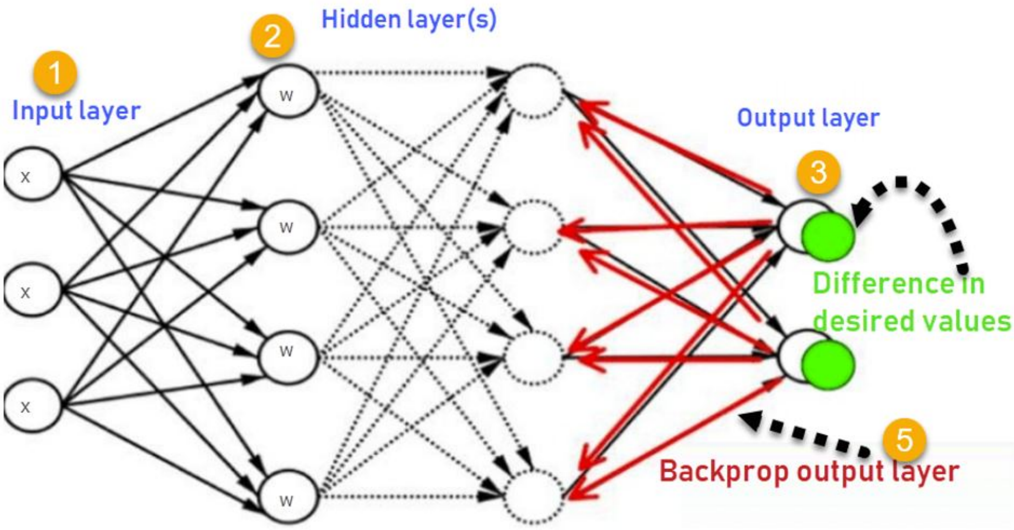

Figure 3.7: Dropout Layer.

1. Inputs X, arrive through the preconnected path 2. Input is modeled using real weights W. The weights are usually randomly selected. 3. Calculate the output for every neuron from the input layer, to the hidden layers, to the output layer. 4. Calculate the error in the outputs 5. To change the weights such that the error is reduced, go back from the output layer to the hidden layer. 6. Repeat the procedure until the desired results are obtained. [14]

Advantages of Backpropagation

Most prominent advantages of Backpropagation are: 1. It's quick, easy, and simple to program backpropagation. 2. Besides the input numbers, it has no parameters to adjust. 3. Due to the lack of prerequisite network expertise, it is a flexible method. 4. It is a tried-and-true approach that typically yields positive results. 5. The characteristics of the function do not need to be mentioned in any particular detail.

Backpropagation Key Points

1. By including weighted linkages that have the least impact on the training network, the network structure is made simpler. 2. To establish the connection between the input and hidden unit layers, you must research a set of input and activation values. 3. Analyzing the effect a specific input variable has on a network output is helpful. Rules should reflect the knowledge acquired from this analysis. 4. Deep neural networks working on error-prone tasks like speech or picture recognition benefit particularly from backpropagation. 5. Backpropagation makes use of the chain, and the power rules enable it to work with any quantity of outputs.

3.16 Padding

Since padding refers to the number of pixels that are added to an image during processing by the CNN kernel, convolutional neural networks (CNNs) can profit from it. For instance, any additional pixels will have a value of 0 if the padding in a CNN is set to zero. The image will have an additional one-pixel border with a pixel value of zero added to it if the zero padding is set to 1, though. [15]

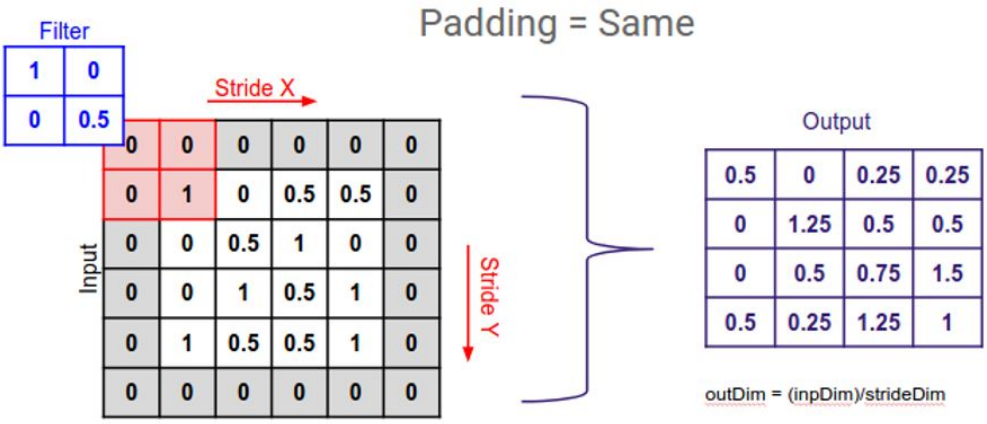

Figure 3.8: Padding Layer.

How does Padding work? Convolutional neural networks can process a larger area of an image when padding is used. The kernel is the neural network filter that goes through the image, scanning each pixel and converts the data into a smaller, or occasionally bigger, format. Padding is added to the image's frame in order to provide the kernel additional room to process the image and aid in its processing. In order to improve the accuracy of image analysis, padding should be added to images processed by CNN.

3.17 13 Non-Linearity Layers

Non-linearity layers are frequently included right after the convolutional layer to add non-linearity to the activation map because convolution is a linear operation and images are anything but linear. Non-linear operations come in a variety of forms, the most common ones being:

1. Sigmoid The mathematical formula for the sigmoid nonlinearity is () = 1/(1+e). It "squashes" a real-valued number into the range between 0 and 1. The gradient of a sigmoid becomes almost zero when the activation is at either tail, which is a very unfavorable sigmoid feature. Backpropagation will effectively "kill" the gradient if the local gradient gets too small. Additionally, if the input to the neuron is exclusively positive, the output of the sigmoid will either be exclusively positive or exclusively negative, leading to a zigzag dynamic of gradient updates for weight. 2. Tanh A real-valued number is condensed by Tanh to the range [-1, 1]. Similar to sigmoid neurons, the activation saturates, but unlike them, its output is zero-centered. 3. ReLU The Rectified Linear Unit (ReLU)

has gained a lot of popularity recently. The function ()=max (0,) is computed. In other words, at 0, the activation is only the threshold. ReLU is more dependable and speeds up convergence by six times when compared to sigmoid and tanh. ReLU's potential for fragility during training is, however, a drawback. A huge gradient passing through it has the potential to update it in such a way that the neuron is never updated further. However, by choosing an appropriate learning rate, we can make this work. Now that we are aware of the different parts, we can construct a convolutional neural network. 60,000 training examples and 10,000 test examples make up the Fashion-MNIST dataset of Zalando article photos that we'll be using. Each illustration is a 28x28 grayscale graphic paired with a label drawn from one of ten classes. The architecture of our convolutional neural network is as follows:

[INPUT] → [CONV 1] → [BATCH NORM] → [ReLU] → [POOL 1] → [CONV 2] → [BATCH NORM] → [ReLU] → [POOL 2] → [FC LAYER] → [RESULT]

We'll use a spatial kernel of size 5 x 5, stride size 1, and padding of 2, for both convolutional layers. We'll utilize the max pool operation with kernel size 2, stride 2, and no padding for both pooling layers.

3.18 Advantages of CNN Architecture

The benefits of a convolutional neural network include the following:

1. Computationally, CNN is effective.
2. Utilizes unique convolution and pooling algorithms in addition to parameter sharing. Now that CNN models can operate on any platform, they are appealing on a worldwide scale.
3. Without the assistance of a human, it finds the necessary features.
4. It can be used in a range of businesses to carry out important activities including, among others, object identification, document analysis, facial recognition, and climate comprehension.
5. You may extract useful features from a trained CNN by giving it your data at each level and adjusting the CNN slightly for a particular task. [16]

3.19 Applications

Convolutional neural networks are employed in the following ways today:

1. Object detection: Thanks to CNN, we now have advanced models like R-CNN, Fast R-CNN, and Faster R-CNN that are the main pipeline for many object identification models used in autonomous vehicles, facial detection, and other applications.

2. Semantic segmentation: In 2015, a team of academics from Hong Kong created a CNN-based Deep Parsing Network to include rich data in an image segmentation

model. Additionally, UC Berkeley researchers created fully convolutional networks that enhanced the most recent semantic segmentation techniques.

3. Image captioning: Recurrent neural networks are combined with CNNs to provide captions for pictures and videos. This can be applied to a variety of tasks, such as identifying activities or providing descriptions of films and images for visually impaired people. In order to make sense of the enormous volume of videos that are routinely posted to the network, YouTube has heavily utilized it. [17]

3.20 Motivation behind Convolution

Convolution makes use of the three key concepts of sparse interaction, parameter sharing, and equivariant representation that drove computer vision researchers. Let's go into more depth about each of them. Matrix multiplication is used in trivial neural network layers to describe the interaction between the input and output unit through a matrix of parameters. This implies that every input unit and output unit communicate with one another. Convolution neural networks, however, only interact sparsely. This is accomplished by making the kernel smaller than the input; for instance, an image may have millions or thousands of pixels, but by utilizing the kernel to process it, we can find significant information that is only of tens or hundreds of pixels. This indicates that we need to keep fewer parameters, which lowers the model's memory need and boosts the model's statistical effectiveness. If calculating a certain property at a given spatial location (x1, y1) is beneficial, then it should also be useful at another location, such as (x2, y2). It indicates that neurons are required to use the same set of weights when constructing an activation map for a single two-dimensional slice, or for a single activation map. A convolution network contains shared parameters, meaning that weights applied to one input are the same as weights applied elsewhere for getting output, unlike a standard neural network where each member of the weight matrix is utilized just once and never again. The layers of a convolution neural network will have the attribute of equivariance to translation as a result of parameter sharing. It asserts that if we adjust the input in a certain way, the output will follow suit.

Chapter IV

Results and Analysis

4.1 Deep Learning

Artificial intelligence (AI) refers to the ability of machines to carry out tasks that ordinarily require human intelligence. It incorporates machine learning, where tools may learn things on their own and improve their abilities. Deep learning, a kind of machine learning, uses artificial neural networks—algorithms based after the human brain—to learn from enormous volumes of data. Similar to how we learn from experience, the deep learning algorithm would do a task repeatedly while making small tweaks each time to improve the results. We refer to this process as ”deep learning” since neural networks have a number of (deep) layers that aid in learning. Almost every problem that requires ”thinking” to solve may be learned to be solved using deep learning. Deep learning helps machines to resolve challenging problems even with a set of data that is extremely different, unstructured, and interrelated. The more they learn, the more effectively deep learning algorithms work.

4.2 Dataset

Chest X-ray analysis is a difficult work in the realm of health to take into account. There are now hundreds of readily available datasets for chest X-rays, however they are only capable of holding a few thousand images. Using chest X-ray pictures and artificial intelligence, the Radiology Society of North America (RSNA) launched a pneumonia detection competition in 2018. (AI). Participants are tasked with developing an algorithm that can identify and classify pneumonia from chest x-ray images. This collection contained pictures of typical chest X-rays with no lung infection and non-COVID pneumonia. The enormously well-liked Kaggle chest X-ray database contains 6432 chest X-ray images of healthy or normal, viral, and bacterial pneumonia, with resolutions ranging from 800 pixels to 1900 pixels. 4303 photos are afflicted by bacterial pneumonia and 1583 images are from healthy or normal chest X-rays out of a total of 6232image datasets. Images of the COVID-19 virus, both positive and suspect, were obtained from publically accessible sources . The most recent database collection was created using chest X-ray pictures from both healthy and pneumonia-affected patients. The images from the database collection of chest X-rays for normal X-ray, COVID-19 infected and Pneumonia datasets are all downloaded through Kaggle.

4.3 Data Process

The RGB pixel formatted patches are all sized from 0 to 224. We intended to use machine learning (ML) categorization techniques to these images. As a result, to be consistent with the methods, we developed a scale that runs from 0 to 1.

4.4 Data augmentation

Data augmentation is an important part of deep learning. In deep learning algorithm we train a lot of data for increasing our accuracy in the output. For this we want to train our network with more and more data. But problem is when we have the limitations of data that time data augmentation helps us a lot. By this we generate more data with our existing data. Suppose we have almost 6000 of images. After using the augmentation our number of image is increased by double, lets take that our present image number is around 1200. How is that works? The algorithm take one image and rotate, blur or crop the image and create a new one. By this process we have a decent size of data set to train the network. [18]

4.5 Image Patch Sampling

The foundation is discarded in patches if it has a lot of greasy tissue or a few slides. A pathologist manually notifies IDC. For training purposes, it is used to build a binary annotation mask. If 80 percent of the data are present in the image patch annotation mask, a positive sample is recommended. A negative sample goes under another name. Each WSI divides the network analysis into non-overlapping 100x100 pixel picture fixes. Each picture patch is converted from RGB (red, green, and blue) to YUV colour, and then normalized to have a mean value of zero and a variance value of one. During this step, you can focus on traits like sparseness that are unaffected by covariance by removing raw pixel correlations.

4.6 Logistic Regression (LR)

The technique of forecasting an outcome, such as whether something will occur or not, is known as logistic regression can be shown, for instance, as ”Yes/No,” ”Pass/Fail,” etc. , which may be defined as, P(Y = 1—X) or P(Y = 0—X) X is an independent variable, then The dependent variable Y is calculated. The logistical regression may be linear or nonlinear with a high polynomial order. [18]

4.7 Performance Measure

Each picture patch was categorized as either positive or negative using the labels 1 and 0. We estimated the accuracy (Pr), recall (Re), and false negatives (FN) by computing the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values (FN). These values were required for the algorithms' performance evaluation. The key performance measurements, or F1 score, are created using an arithmetic and harmonic mean.

$$Accuracy = \frac{TP + TN}{P + N} \tag{4.1}$$

$$Precision = \frac{TP}{TP + FP} \tag{4.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{4.3}$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4.4}$$

Where,
TP: correct positive prediction,
FP: incorrect positive prediction,
TN: correct negative prediction,
FN: incorrect negative prediction,
P: TP+FP,
N: TN+FN.

4.8 Methodology

We must first perform a straightforward image processing procedure before beginning our CNN architecture method. Therefore, we begin by gathering our image dataset. We must examine our dataset. We are now using the Jupiter notebook program. We need to have a solid understanding of Python before we can begin this project. We load our data photos first. Even though there are more than 6,000 photographs in this dataset, it takes a while to finish. We must change the image labelling. It separated our image into the pixel intensities of each. The image pixel size in this dataset is 224x224x3. Images must occasionally be scaled because they are too large. We don't need to resize our data photos in this situation. Next, we prepare our dataset. Our RGB pixel intensity histogram images are plotted. The data's size ranges from 0 to 224, but its scale is between 0 and 1. As a result, a variety of strategies can be used to categorize the data. One last change was made to the Image Data Generator Keras function before training. The objective is to increase the amount of training data by only altering the current data—using shifts, rotations, brightness changes, horizontal/vertical inversions, etc. Only after separating the data from training and testing is this function used. A training dataset and a testing dataset were created using our dataset. We use a 80 percent training dataset and 20 percent testing dataset to avoid overfitting or bias. How rapidly the weights in the neural network change depends on the learning rate. The algorithm's epoch number indicates how many times the data is displayed, and each new period also updates the weight value. Every training cycle uses a certain amount of data, which is sample size. The weights updating algorithm is indicated by the optimizer value.

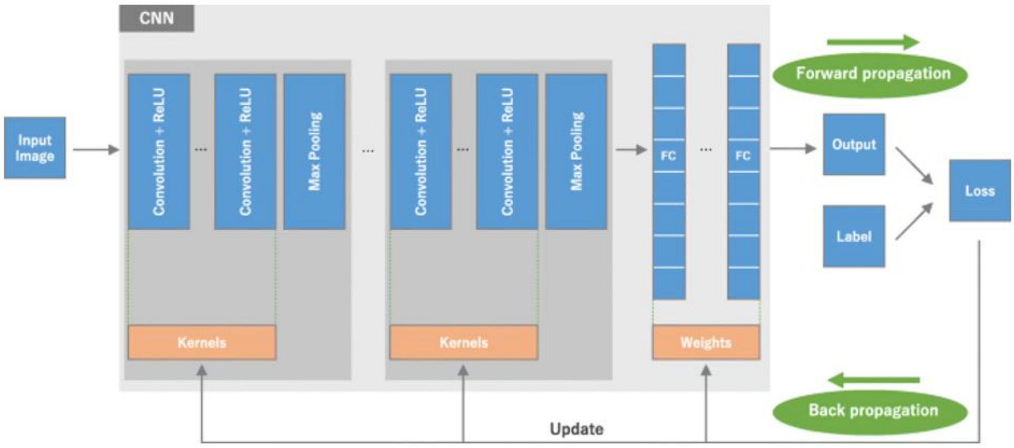

Figure 4.1: CNN Architecture.

4.9 Keras

When neural networks were first developed, programmers who were not familiar with deep learning were unable to use the principal systems because of their depth. There are a number of suggestions for improved and condensed high-level APIs for creating neural network models, many of which at first glance seem to be similar but differ upon closer inspection. New architecture can be combined with separate components such neural layers, optimizers, activation functions, and regularization methods. New modules and techniques can be added with ease. Model specifications are made in python code as opposed to separate model configuration files. Python's Keras framework enables the creation and evaluation of deeper learning models.Both powerful and simple to use. The most used APIs are those for CNN. It's the dialect of Python.To assess the DN model, several open sources are available. With just a few lines of code, you can build and train neural network models using Theano and TensorFlow, two effective numerical computation frameworks. [28] Both programs enable the training of neural networks on either a CPU or a GPU, the latter being more important due to its shorter training time Because Keras by itself does not support low-level operations like tensor and convolution multiplication, TensorFlow will be used as a backend tool in this project.

4.10 Experimental Result

In this project we are trying to classify an image data. By this thing we almost work on more than 6000 of image data. In this project we are classifying a image data in between three category. The categories are COVID-19, PNEUMONIA and NORMAL. For this we train our project with 80percent of our image data and test our project with 20percent of datas. We use a CNN model for classify our image. The architecture of CNN is in below

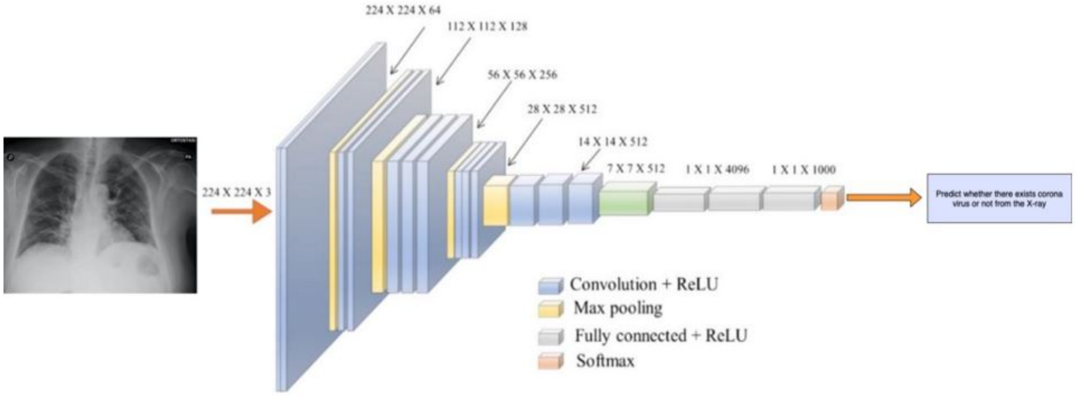

Figure 4.2: CNN Architecture for Covid-19 Pneumoia Detection.

In this figure we can see the architecture of CNN model. We can see the feature map size and the other things what happens inside the model. Now lets see the flow chart for this network

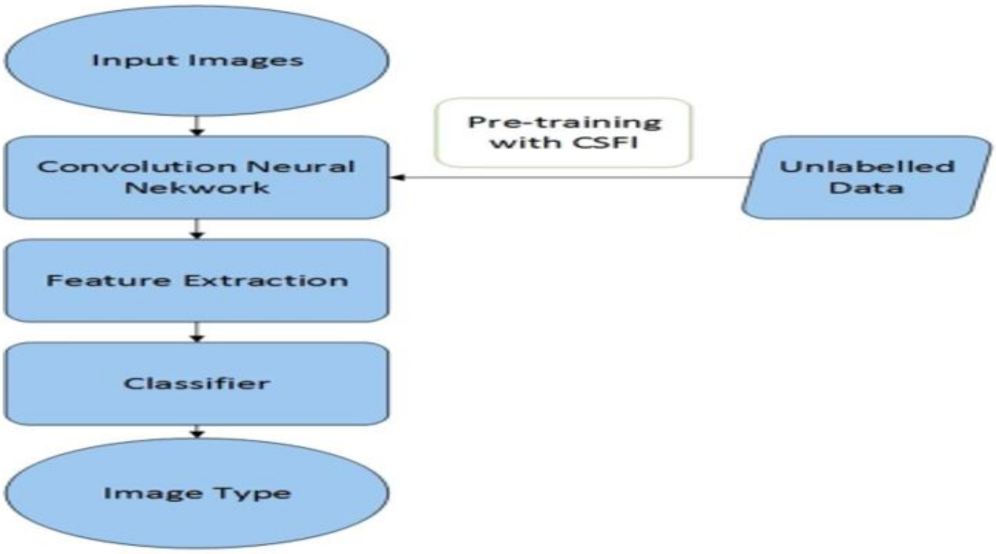

Figure 4.3: CNN Architecture Flowchart.

Here we created our own CNN architecture which is inspired by VGG16 architecture. Here we are using 4 convolutional layer , 4 maxpooling layer, 3 dropout layer , and 2 dense layer in this architecture. In this architecture we are using 128 steps per epoch and total epoch size is 111. By this we are trying to make sure our accuracy must be around 95 to 100 percent. And we succeeded in our destination. We have more than 95 percent of accuracy. Which is pretty good as a new architecture.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 64)	1792
max_pooling2d (MaxPooling2D)	(None, 111, 111, 64)	0
conv2d_1 (Conv2D)	(None, 109, 109, 128)	73856
max_pooling2d_1 (MaxPooling2	(None, 54, 54, 128)	0
dropout (Dropout)	(None, 54, 54, 128)	0
conv2d_2 (Conv2D)	(None, 52, 52, 256)	295168
max_pooling2d_2 (MaxPooling2	(None, 26, 26, 256)	0
dropout_1 (Dropout)	(None, 26, 26, 256)	0
conv2d_3 (Conv2D)	(None, 24, 24, 512)	1180160
max_pooling2d_3 (MaxPooling2	(None, 12, 12, 512)	0
dropout_2 (Dropout)	(None, 12, 12, 512)	0
flatten (Flatten)	(None, 73728)	0
dense (Dense)	(None, 512)	37749248
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 3)	1539

Figure 4.4: Used CNN Architecture for Covid-19 Detection.

By using this architecture we have a loss function. Which we denoted by loss and validation loss. Lets check the loss function graph.

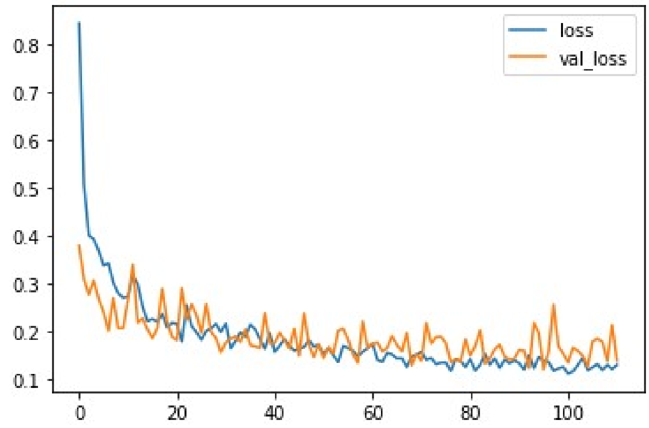

Figure 4.5: Used CNN Architectures Validation Loss.

We have a good accuracy from our architecture. which is more than 95 percent. Let's check that one.

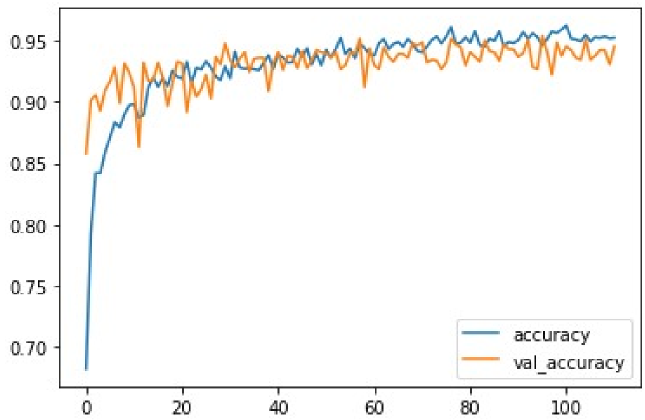

Figure 4.6: Used CNN Architectures Validation Accuracy.

We have a matric result of our architecture . this matric gives us a view that how our architecture works.

	precision	recall	f1-score	support
0	0.99	0.96	0.97	116
1	0.89	0.91	0.90	317
2	0.96	0.96	0.96	855

Figure 4.7: Matric Result of CNN.

We try to figure out the accuracy and loss together in a graph. For this we plot a graph of accuracy vs loss. The graph is given below

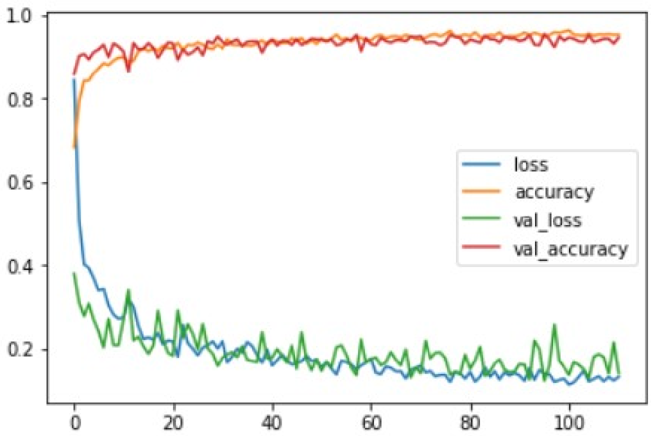

Figure 4.8: Validation Loss VS Validation Accuracy.

We have a test dataset which has 3 classes. The classes are COVID19,PNEUMONIA and NORMAL. We plot a matrix so that we have a visual representation of accuracy output. The matrix is called as Confusion matrix. The matrix is in below

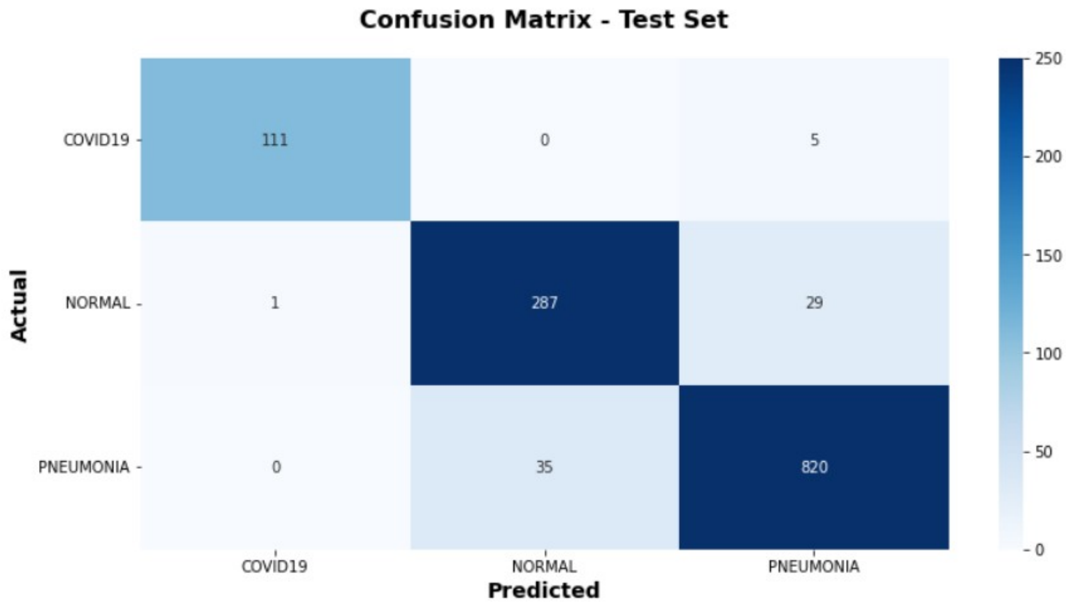

Figure 4.9: Confusion Matrix of this Architecture.

All of this output came from our architecture which is based on CNN. We want to visualize the parameters and other things step by step though a figure. By this figure we will able to understand the working process of our architecture. This architecture gives us a good prediction which we want to have. In this architecture we want to achieve the accuracy of 98 to 100 percent. But our architecture is not able to achieve that.in this architecture we will able to see the all of the layers one after another is positioned. by this figure, we will able to understand about the hidden work of thew network. How the parameters are changing, which layer comes after which layer. This figure tells us the summary of the working process of the network.the figure is given below:

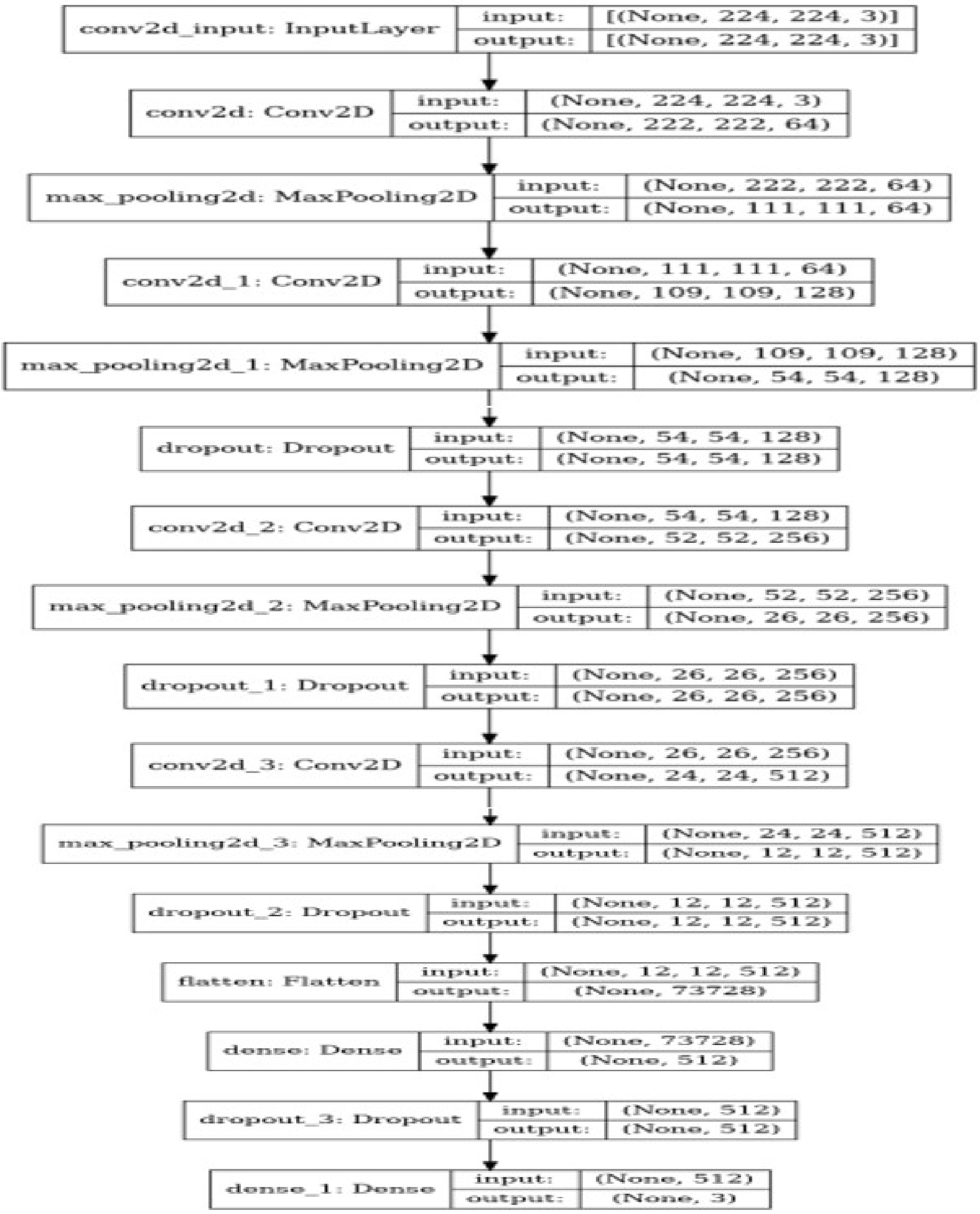

Figure 4.10: Layer By Layer Of this CNN Architecture.

Chapter V

Conclusion and Discussion

5.1 Conclusions

In this CNN architecture we are trying to get the best out put as possible. In this case we try almost 6000 image data to complete our prediction. This prediction depends on our epoch size, learning rate, dataset, the layers of network and also on activation function. After that we have a more than 95 percent of accuracy. We tried to make the best use of our data. We can do the augmentation to increase our data but we are willing to use the raw data of our dataset. For this we have a dataset which is not sufficient for us. If we are able to have a good amount of data and we use the activation function and learning rate good and standard size of epoch then we will able to increase the accuracy. There are a lot of pre-built architectures are there but we choose to accept the challange to built a new one and we succeeded in our work. We named it ”Sequential”. as this architecture gives us a very good amount of accuracy as our first trial. We are very happy with this architecture.

5.2 Future Directions

This study proposes a two-stage deep residual learning method to detect pneumonia caused by COVID-19 using lung X-ray pictures. Using the VGG16 model, the model performed well in identifying COVID-19 patients from those who had COVID-19-induced pneumonia. With an average sensitivity of 95.92 percent , specificity of 100 percent, and accuracy of 91.69 percent, the model accurately predicted pneumonia. Accuracy is raised while training loss is decreased. In the current situation, parallel testing can be utilized to stop the illness from spreading to front-line personnel and to produce first diagnostics that show whether a patient has COVID-19. As a result, the suggested technique can be utilized as a substitute diagnostic tool to identify pneumonia cases. By modifying the hyperparameters and transfer learning combinations, future research can enhance the CNN architecture’s performance. An upgraded, complicated network structure might be another practical technique to choose the optimal model for pneumonia and COVID-19. We also want to develop our architecture one or two or three step more. so that we can able to ensure a better Accuracy and precision in this architecture. The work is not all over. this is the primary level of work.We want to Go farther More with this Project.

Bibliography

[1] P. R. Bassi and R. Attux, "A deep convolutional neural network for covid-19 detection using chest x-rays," *Research on Biomedical Engineering*, vol. 38, no. 1, pp. 139–148, 2022.

[2] I. Fahmi, "World health organization coronavirus disease 2019 (covid-19) situation report," *DroneEmprit*, 2019.

[3] S. Akter, F. J. M. Shamrat, S. Chakraborty, A. Karim, and S. Azam, "Covid-19 detection using deep learning algorithm on chest x-ray images," *Biology*, vol. 10, no. 11, p. 1174, 2021.

[4] A. Alhudhaif, K. Polat, and O. Karaman, "Determination of covid-19 pneumonia based on generalized convolutional neural network model from chest x-ray images," *Expert Systems with Applications*, vol. 180, p. 115141, 2021.

[5] C. N. Rachmi, K. E. Agho, M. Li, and L. A. Baur, "Stunting, underweight and overweight in children aged 2.0–4.9 years in indonesia: prevalence trends and associated risk factors," *PloS one*, vol. 11, no. 5, p. e0154756, 2016.

[6] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Computational and structural biotechnology journal*, vol. 16, pp. 34–42, 2018.

[7] A. Gulli and S. Pal, *Deep learning with Keras.* Packt Publishing Ltd, 2017.

[8] uknown, "convolution layer," 2021.

[9] W. Wu, S. Ma, Y. Su, and C.-H. Wu, "Double-layer learning, leaders' forgetting, and knowledge performance in online work community organizations," *Journal of Organizational and End User Computing (JOEUC)*, vol. 33, no. 1, pp. 92–117, 2021.

[10] uknown, "How relu and dropout layers work in cnns," 2022.

[11] A. Roy, "An introduction to gradient descent and backpropagation," *source: https://towardsdatascience. com/an-introduction-to-gradient-descent-and-backpropagation-81648bdb19b2*, 2020.

[12] J. Brownlee, "Gentle introduction to the adam optimization algorithm for deep learning," *Machine Learning Mastery*, vol. 3, 2017.

[13] D. Johnson, "Back propagation in neural network: Machine learning algorithm.," *source:https://www.guru99.com/backpropogation-neural-network.html*.

[14]

[15] Unknown, "What is padding in machine learning?," *source:https://deepai.org/machine-learning-glossary-and-terms/padding*.

[16]

[17] M. Mishra, "Convolutional neural networks, explained," *source:https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939*.

[18] S. A. Alanazi, M. Kamruzzaman, M. N. Islam Sarker, M. Alruwaili, Y. Alhwaiti, N. Alshammari, and M. H. Siddiqi, "Boosting breast cancer detection using convolutional neural network," *Journal of Healthcare Engineering*, vol. 2021, 2021.