Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

Empirical Evaluation of Pre-Trained Deep Learning Networks for Pneumonia Detection

Shahab Uddin Agha[1], Dr. Muhammad Shahid[2], Farhan Mansoor[3], Sumaiya[4]
[1,2,3,4]Department of Computer Science, Alhamd Islamic University, Quetta.
shahab.agha.it@gmail.com

DOI: 10.5281/zenodo.10185080

ABSTRACT

Pneumonia is a significant global health issue, characterized by a substantial mortality risk, impacting a vast number of individuals on a global scale. The quick and precise identification of pneumonia is crucial for the optimal treatment and management of this condition. This research work aims to answer the pressing need for precise diagnostic methods by using two advanced deep learning models, namely VGG19 and ResNet50, for the purpose of pneumonia detection in chest X-ray pictures. Furthermore, the present area of research is on the use of deep learning methodologies in the domain of medical image analysis, namely in the identification of pneumonia cases via the examination of chest X-ray images. The study challenge pertains to the pressing need for accurate and automated pneumonia diagnosis to assist healthcare professionals in making timely and educated judgements. The VGG19 and ResNet50 models were trained and assessed using the comprehensive RSNA Pneumonia dataset. In order to enhance their performance in the classification of chest X-ray pictures as either normal or pneumonia-affected, the models underwent rigorous training and meticulous fine-tuning. Based on the results obtained from our investigation, it was seen that the VGG19 model exhibited a notable accuracy rate of 93\%, surpassing the ResNet50 model's accuracy of 84\%. Furthermore, it is worth noting that both models demonstrated a notable level of precision, recall, and f1-scores in the identification of normal and pneumonia patients. This indicates their potential for accurately classifying such instances. Furthermore, our research findings indicate that deep learning models, namely VGG19, have a high level of efficacy in reliably detecting pneumonia via the analysis of chest X-ray pictures. These models has the capacity to function as helpful tools for expediting and ensuring the precise identification of pneumonia by healthcare practitioners.

Keywords: Machine learning, Deep learning, Transfer learning, ResNet, VGG19, AI, CNN, Pneumonia, Radio graphy.

Cite as: Shahab Uddin Agha, Dr. Muhammad Shahid, Farhan Mansoor, & Sumaiya. (2023). Empirical Evaluation of Pre-Trained Deep Learning Networks for Pneumonia Detection. *LC International Journal of STEM*, *4*(3), 47–81. https://doi.org/10.5281/zenodo.10185080

INTRODUCTION

Pneumonia is a potentially dangerous lung illness that damages the air sacs and causes symptoms such as chills, fever, coughing up mucus, and trouble breathing. It may be caused by bacterial, fungal, or viral infections, resulting in fluid buildup in the air sacs. This illness may spread fast in patients with compromised immune systems, including infants and older children with impaired immune systems.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

Sadly, pneumonia remains a big hazard, with over one million children dying globally in 2018. (Rui, Kang, & Ashman, 2018).
Medical professionals may diagnose pneumonia using imaging techniques such as radiographs, computed tomography scans, or magnetic resonance imaging. In order to determine whether or not a patient has pneumonia, an evaluation of their chest x-ray using the methods mentioned earlier may be performed. Confirm the patient's diagnosis by going through their medical history and the results of their tests. Diagnosing and treating pneumonia as soon as possible is necessary to forestall the development of major complications or fatal outcomes.

If an X-ray is allowed to pass through the chest, the resulting radiograph will show the chest's soft tissues as black and its hard tissues, such as the bones, as white. Those persistently ill and have signs of fluid buildup in the lungs and chest cavity are said to have pneumonia. Lung air sacs may be identified on a chest x-ray by their lighter hue, a telltale sign of their presence. It is possible that the patient has several lung abnormalities, such as covid-19, cancer cells, blood cells, other red cells, heart abnormalities, and dilated blood vessels. In most cases, a chest x-ray is used to confirm the existence of a lung infection's existence and determine its severity and localization within the lung. Optimal productivity. Therefore, the project is successful in terms of efficiency benefits in medical emergency processing in remote areas where pneumonia and covid-19 is prevalent. CNN model training and assessment for classification purposes was possible for scientists. Imaging the chest for normal and infected conditions uses several different classifiers.

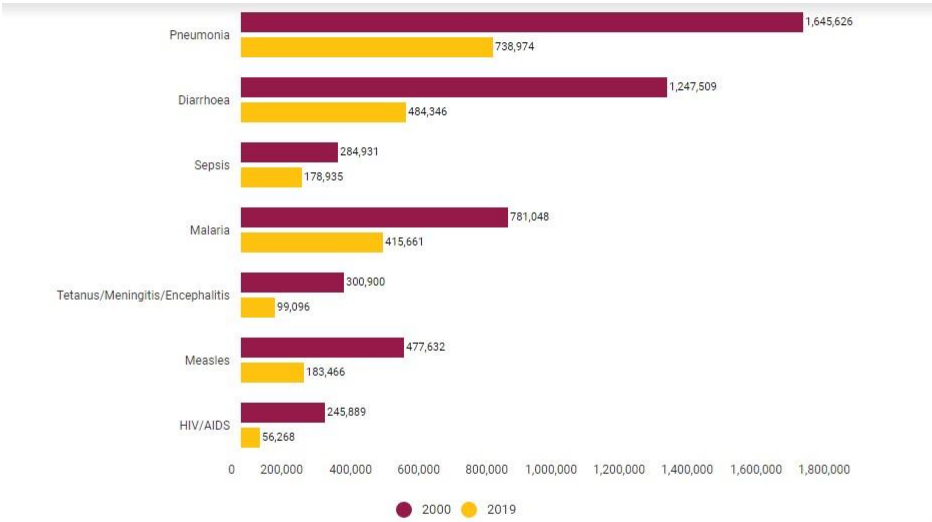
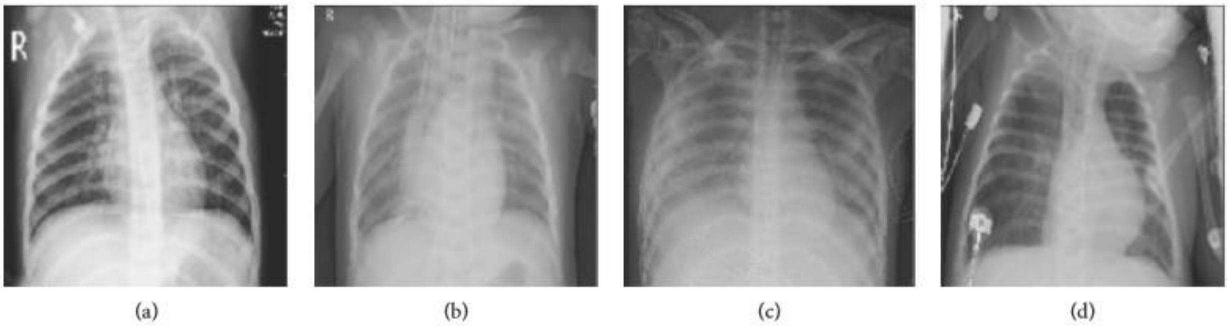


Figure 1.1: Comparing the rates of infant mortality due to infectious diseases in the years 2000 and 2019 (Rui et al., 2018)



Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

Figure 1.2: X-rays images pneumonia effected individual (Rui et al., 2018)

Recent improvements to Computer-Aided Design (CAD) software have made this an important goal for AI and ML researchers. CAD systems have helped doctors find breast cancer, classify diseases using mammography, find lung cancer, and do many other things. Medical imaging CAD systems are becoming useful tools for figuring out what is wrong with someone and how bad it is. On the way to a correct diagnosis, the medical team uses CAD to help them make decisions and ensure they are right. When trying to make anything brighter, it is preferable to use approaches from machine learning rather than handmade features since handcrafted features cannot extract all of the useful information from a picture (Poostchi, Silamut, Maude, Jaeger, & Thoma, 2018). The future will be good for the next generation. Since its start, deep learning has come a long way in simulating how the human brain works. In particular, it gives people a way to deal with their real problems. Convolutional neural networks (CNNs) enabled by deep learning can effectively analyze medical images by extracting essential features for classification, leading to promising results (Fafi, n.d.). CNNs have the advantage of providing accurate classification probabilities with a small amount of input data (Abiyev & Ma'aitaH, 2018). Recent research has focused on improving the performance of deep-learning models to make CNNs better at spotting pneumonia.

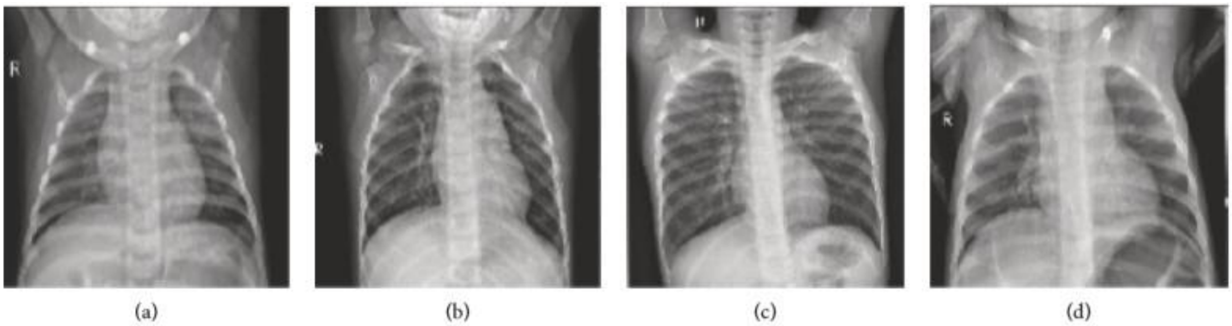


Figure 1.3: X-rays images of normal individual (Stephen et al., 2019)

DEEP LEARNING PHENOMENON

Have you ever wondered how Google can instantaneously translate an entire web page or how your phone can organize photos according to where they were taken? Deep learning is to blame for everything that has happened here. As seen in the graphic titled artificial intelligence (AI) is a method that enables robots to behave in a manner that is analogous to that of humans. Deep learning is a strategy that takes its cues from the workings of the human brain. In contrast, machine learning is more concerned with tackling any given problem by combining several different algorithms.

ARTIFICIAL NEURAL NETWORK

The method of detecting alien or harmful cell incursion in human body cells is known as the malaria illness diagnosis process. Through various methods and approaches, these undesirable inductions must be visualized or recognized. In these processes, biomedical pictures are essential. These images depict not only the undesired parasites or cells but also all of the body's inner organs and tissues.

Today's computer-assisted tools and methodologies allow for a more accurate evaluation of these visual representations. This innovation has gotten a lot of help from machine learning. Learning the patterns in these visual images is part of the supervised machine learning process, as is detecting and creating a mathematical formula to locate the corresponding pattern in the same photos. This entire process is referred to as training images with machine learning and generating a mathematical formula that can be applied to the same type of image for future classifications. The deeper the network, the more precise the patterns discovered, meaning that more computing and time resources are required.

For classification, machine learning approaches are separated into two categories. Learning is divided into two categories: supervised and unsupervised (Xue, Yuan, Wu, Zhang, & Liu, 2020). A model will be provided both parasite and nonparasite images, as well as a target class classifying the image as a parasite or nonparasitic, during supervised learning. where only photos are produced by unsupervised learning We will use supervised learning techniques because our model is in a binary target class. Traditional algorithms for pattern recognition and classification include Decision Tree, Naive Bayes, and SVM. When it comes to large datasets with images, all of these models operate better for tiny and numerical data predictions. To match our categorization accuracy, we need to do more in-depth analyzing and searching. By using algorithms that can learn from large datasets without explicit programming, the field of deep learning seeks to overcome constraints and fill in gaps (Goodfellow, Bengio, & Courville, 2016). Furthermore, conventional algorithms were all flat. Flat in the sense that it requires feature extraction and pre-processing, but deep neural networks may utilize all types of data.

NEURAL NETWORKS

By their name, neural networks imply that they are modeled after the human brain’s neural system, in which every neuron is linked. This entire network of neurons collaborates to process any type of data. As a result, it’s sometimes referred to as an Artificial Neural Network.

An input layer, a middle layer with some functionality concealed from view, and an output layer make up this structure. Each of these layers is composed of neural nodes, which may be compared to mathematical formulae if you prefer that analogy. These neural nodes make up the core of this network and are an integral part of it. Each node or neuron in this network has its own set of parameters (which we’ll go over later) and functions. Now These layers must be traversed to find some pattern or linkages among a collection of data. This procedure will be referred to as training. It could be made up of numerous passes. We identify the mathematical relationship in input data by repeatedly going through these levels. This process is known as data training. Once these models have been properly trained, we are ready to classify or forecast fresh data. In a nutshell, any neural network can learn any mathematical function with a small amount of training.

1. Artificial neural networks, also known as ANNs, are made up of numerous layers of linked nodes referred to as neurons.
2. The functionality of each artificial neuron is modeled after that of a real neuron found in the human brain.
3. The first layer present in an ANN is referred to as the input layer, followed by one or more hidden layers and, finally, the output layer.
4. A calculation is carried out by each node within the network, and the results of that computation are then passed on to succeeding nodes, which results in a chain reaction of information flowing across the network.

DEEP LEARNING

In the realm of biological detection and classification, the revolutionary technique of deep learning has caught the attention of researchers worldwide, as revealed by the groundbreaking study by Mahmud and colleagues (2018). This avant-garde method features a dynamic middle layer composed of multiple neurons, allowing for unparalleled accuracy and precision in the detection and classification of biological data. The secret to this success lies in the intricate network of hidden layers designed to identify complex patterns in enormous data sets, leading to astute predictions. Deep learning models are distinguished by their innumerable hidden layers, far surpassing the typical neural network's meager two or three hidden layers. However, an extensive amount of labeled data is a prerequisite to training these models. The data labeling process entails identifying the target variable, either for the input data or the image output, enabling the prediction of highly accurate outcomes.

SINGLE NEURON UNIT

A neural network is made up of nodes (neuron connections) that are linked in some way. Every input is coupled by a number called weights, and each neuron is connected by a number called bias value (Gonzalez-Abril, Angulo, Velasco, & Ortega, 2008). In a neural deep network, all connected units work in the same way. A mathematical function is called a neuron (unit). Taking single or numerous inputs and using equations to generate an output.

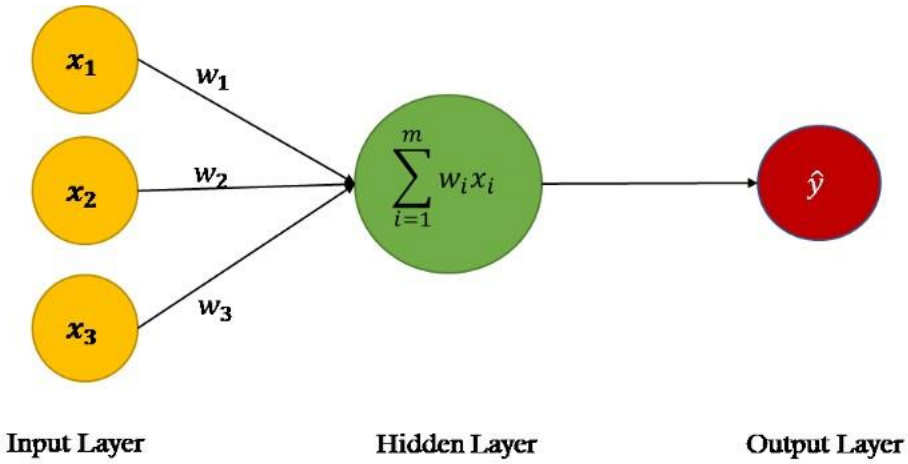


Figure 1.4: Basic neuron (unit) working (Gonzalez-Abril et al., 2008)

- The input data consists of x1, x2, and x3.
- Each input is associated with a weight, represented by w1, w2, and w3.
- The processed output function is referred to as the activation function, denoted by y.

FORWARD PROPAGATION

In a neural network, forward propagation refers to a whole cycle of input data going through a deep neural network model and creating output, followed by a comparison of the projected output to the actual output to determine the inaccuracy.

Steps include:

- The input data is passed into the neurons of the first layer.
- Weights are a connection between each input and the output (any random numerical value). And all of these inputs are summed together at a single-neuron unit after multiplication with their respective weights. In addition, bias was provided as a numerical value.
- Any infinite value can be computed as the output. It is then given to a mathematical function called the Activation Function to limit the output. As a result, any range value is produced.
- This output is subsequently forwarded to a different layer, where it is seen by all of the neurons in that layer. when the same process is repeated with different numbers and functions.
- The output value is generated by the neurons in the final layer. For a given input, this is the expected value.

Activation Functions

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

Activation functions are very important in convolutional neural networks because they decide whether or not an individual neural unit network will transmit its output to subsequent neural layers during a forward pass (You et al., 2020). Choosing the appropriate activation function is crucial for creating an effective neural network. The sigmoid function is commonly used for binary classification, as our data is stored in binary form (Bircanoglu & Arca, 2018; Chandra, 2004). This function appears in the final layer and takes any real value x, determined by neuron weights, and Euler's number e as inputs.

CNN

The use of convolutional neural networks, or CNNs, has become more important in the image recognition, classification, and identification processes (Krizhevsky, Sutskever, & Hinton, 2012). CNNs can learn complicated features using linear convolution, non-linear activation functions, and pooling layers. These networks use the input data as a multi-layer matrix and then apply a smaller kernel matrix to retrieve information pertinent to the problem. CNNs are composed of many basic layers, the first three of which are the convolutional layer, the pooling layer, and the fully connected layer. These components carry out their operations through local connections with common weights, layers, and pooling, followed by the connection of all dense neurons.

CONVOLUTIONAL LAYER

It is the responsibility of the convolutional layer of the CNN architecture to apply a filter to an image and generate feature maps that capture different patterns and features. This layer is an essential part of the architecture of CNN. Image identification and object detection are two important activities that rely on this layer. This approach creates a tiny matrix that may be applied as a feature to the input picture by using a small matrix element-wise multiplication on the matrix to construct the matrix. A stride number is used to provide the step value by which a filter will move in an input image when many filters are utilized for multiple channel images or a grayscale image.

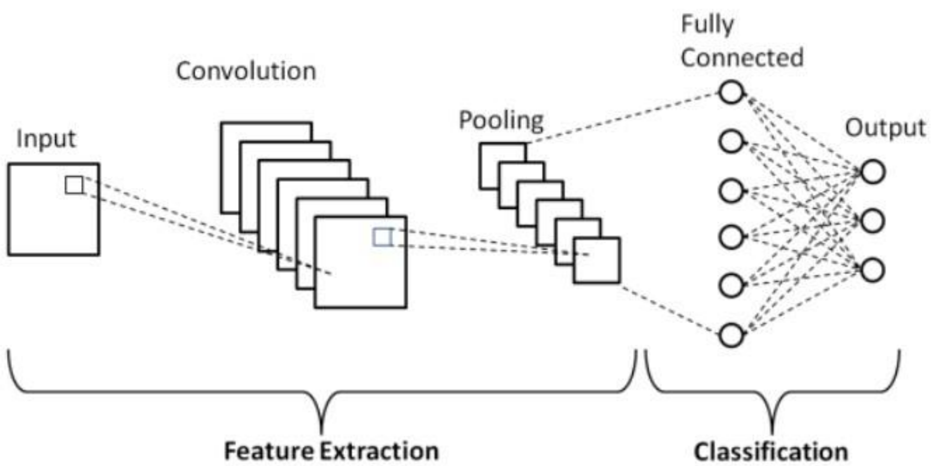


Figure 1.5: Convolutional Neural Network Basic Blocks (Phung & Rhee, 2019)

Pooling

A summary matrix of an extracted feature by filters is used in the pooling operation. Convolutional filters are piled one on top of the other. Its major goal is to cut down on computing time by minimizing the first layer's complicated extracted feature. These new layers make it possible to reduce the size of features. In a nutshell, a pooling layer is a smaller filter that is applied to the filtered image.

Fully Connected

A full-fledged neuronal network. By flattening the data into a vector, all of the relevant characteristics or information retrieved by earlier filters and pooling layers is input into this network. This is placed in the first layer of a dense network, where it creates a mathematical model from which the pattern can be calculated. This dense neural network can have multiple layers, but since we're predicting between infected and uninfected cell pictures, the last layer we'll utilize will only have one neuron. As a result, we can classify things into two groups.

Transfer Learning

When there is a minimal amount of data in a deep learning model, the most prevalent problem occurs. Expanding the amount of data in neural networks can enhance their performance significantly. However, this requires substantial computing power, making it challenging for many neural networks to compete and perform well. The need for computational capacity becomes even higher in the case of biomedical imaging due to image processing.

This issue can be addressed by transfer learning. In general, the knowledge of a previously constructed and trained network or model is applied to a new but related problem. Large datasets are used to train these models. Translation of knowledge The weights and parameters are the only things that are transferred to fresh data. As a result, it is a model that may be fine-tuned to obtain the desired results at its final completely linked layer. The primary advantage of transferring or making use of a pre-trained model is the time and resource savings that may be realized as a result of doing so. The ability to customize the last fully connected layer by tuning these pre-trained models provides flexibility. There are different pre-trained models available with few are:

- VGG16
- ResNet50
- Inception
- MobileNet
- DenseNet

Inception Model

Conventional "filters" can learn linear patches or functions from their input, but M.Lin(Lin, Chen, & Yan, 2014) proposes using multi-layer filters to learn non-linear functions as well. The goal is to converge the outputs of multiple filters into a single vector for another layer. In order to provide evidence in favor of this approach, several models were developed and then trained on the pictures (Deng et al., 2009) dataset. Where a total of 1000 classes must be classified. The objective for using pre-trained models for feature identification is to reduce computing time while also taking advantage of their trained knowledge and applying it to classify our data. Because we just need to categorize between two outputs, we need to update the last layer of the output layer.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

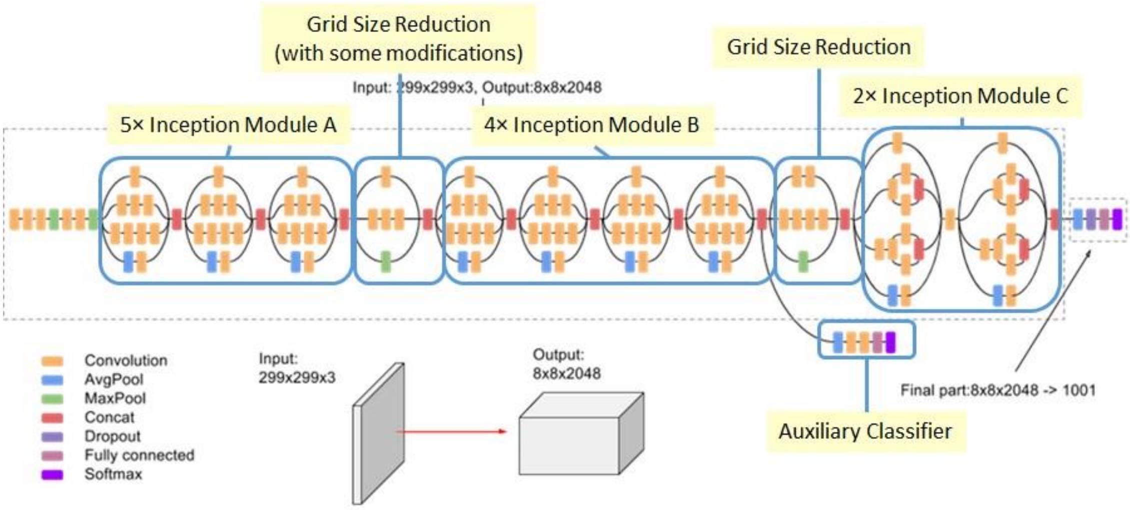


Figure 1.6: Inception module in Inception Architecture

Figure 1.6 demonstrates how the Inception module simultaneously applies several convolution filters to a single input. After that, max pooling is carried out on all the results, which are then concatenated and serve as input for the subsequent Inception module. The 1 * 1 convolution method greatly reduces the parameters. Google (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) has proposed the inception model, often known as GoogLeNet. The Inception model won the ILSVRC2014 ((ILSVRC2014), n.d.) pictures classification competition with an error rate of 6.4 percent, compared to 16.4 percent for the previously trained model. It was a close call in terms of classification accuracy. Although it is based on CNN architecture, there are other features and aspects in the layers, such as the inception module. It is polished in a variety of ways. First, Batch cite ha normalization, also known as inception version 2. Then there are ideas for factorization.

Table 1.1 Inception Model Vs Alexnet Model

Model	Inception	Alexnet
Proposed Year	2014	2012 .
Layers	22	8
Loss error in ILSVRC	6.7%	16.4%
Fully Connected Layers	1	3
Parameters Trained	5975602	20176258

As we can see in Table 1.1 that the inception model to its previously the Alexnet model. It has 14 extra layers with an error rate minimizing till 6.7% compared to the previous 16.4%. And still, it training fewer parameters which means that it’s taking less time compared to the Alexnet model.

LINK BETWEEN DEEP LEARNING, CNN AND INCEPTION MODEL

Deep learning network always performs better when its been provided with clear and small features. Therefor CNN is used to extract that platform for its deep neural calculations. And inception model is the best example of such grafting. Complex mathematical calculations need a heavy processing resource and inception model in this context is a chosen player. That’s why we see a great connection of these three for diagnosis task.

THE DRIVING FORCES BEHIND, AND THE DIFFICULTIES OF, CLASSIFYING LUNG DISEASES

This SECTION addresses the challenges, limitations, and potential currently connected with classifying lung illnesses for therapeutic purposes. A condensed introduction to some fundamental medical imaging and image processing concepts comes first. An explanation of conventional radiographic imaging about the chest X-ray is provided. After that, an analysis of the main factors that triggered a change in the prevailing paradigm of automated image processing is provided to the audience. In addition, this location provides access to the relevant literature and open-source datasets for doing illness analysis using chest X-rays.

Moreover, medical imaging is the non-invasive viewing of interior organs and structures in two or three dimensions. Medical imaging revolutionized medical technology in the 20th century. It helped us understand human anatomy, physiology, and disease patterns. Medical graphics provide clinicians with an objective basis for diagnosing ailments and improving patient care.

Imaging techniques used in medicine may be roughly classified as either projection-based or sectional imaging approaches (Larson & Mostafavi, 2019). While projection imaging has a low cost per test and requires minimal time to acquire data, it is limited to producing flat images. Tomographic imaging, on the other hand, is capable of reconstructing volumetric three-dimensional pictures; however, the process of reconstruction is computationally demanding and needs a significant increase in the amount of time necessary for picture collection. Additionally, medical imaging methods may use either ionizing or non-ionizing radiation. Non-ionizing radiation, such as MRI, US, and MPI, is safe for patients. Ionizing radiation, used in traditional radiography, CT, and PET, can potentially cause mutations in cells, but the benefits of these methods often outweigh the risks. Radiography systems, which are imaging devices that make use of X-rays, are the sort of imaging that is used in clinical practice the most often.

In addition, the fundamental goal of radiology is to retrieve data from images that may be utilized in clinical settings to identify illnesses, do analyses of those illnesses, and manage such illnesses. Radiology also includes surgical intervention, such as the placement of a stent. During this procedure, real-time imaging is utilized to guide radiologists through the body’s organs, blood vessels, and other anatomical structures to reach the body’s internal structures that are being targeted. In recent decades, medical imaging has become a key diagnostic tool and computer resources have grown, enabling medical picture analysis. Medical image analysis aims to help radiologists interpret images. Reproducible, quantifiable, and objective evaluations of medical scans are made easier with these procedures. Analysis of medical pictures is a helpful tool for specialists, who often make qualitative and subjective evaluations of medical images. The field of medical image analysis may be loosely broken down into three primary subfields.

Classifying a new picture involves selecting the appropriate category from a group of available options and assigning that category to the image in question. In medical imaging, one of the most significant tasks is identifying whether or not a pathology is present.

The image registration technique involves aligning two or more photographs to achieve anatomical congruence. Image registration may also refer to the act of registering images. In medical imaging, ct scanning (CT) and positron emission tomography (PET) are two imaging techniques that may assess metabolic information and anatomical aspects. Image registration is necessary to align both scans so that they may be shown together in an overlay.

Image segmentation involves dividing a picture into its constituent parts. Medical imaging segmentation of organs, illnesses, and tissue groupings is crucial. Moreover, this helps determine organ size, shape, and texture in medical images.

Over the last decade, there has been a paradigm change in medical image analysis. This is partly because of the enormous success of deep learning approaches, which can attain super-human performance in various tasks.

Imaging Using Conventional Radiography

When Wilhelm Rontgen developed the first two-dimensional X-ray image of a human body part in 1895, as illustrated in figure 1.7, he achieved a groundbreaking discovery that ushered in a new era in the field of medical imaging (AB, n.d.). Imaging using X-rays has emerged as one of modern medicine's most common diagnostic procedures. In conventional radiography, an item is irradiated with x-rays detected by a detector after being projected onto the object. X-rays can pass through a wide variety of materials, and the density of the substance determines how well they are distributed and how much they are attenuated. Attenuation coefficients in medical imaging may vary greatly due to the different densities of tissues, such as bones, fluids, and other tissues (Bushberg, Seibert, Leidholdt Jr, & Boone,2012). X-rays are converted into pictures using digital X-ray detectors, which are used in modern radiography equipment. A typical digital detector, as stated by (Healthcare, n.d.), has an active image area that is 34.48 with a resolution of 42.12 and can produce an image of 2330 by 2846 pixels with 14bits. Digital detectors can convert direct and indirect radiation, with the indirect conversion being the more common method (Bushberg et al., 2012). Direct conversion detectors convert X-rays directly into electrical signals (Healthcare, n.d.).

Figure 1.7: First ever X-ray image (AB, n.d.)

In modern radiography equipment. A typical digital detector, as stated by (Healthcare, n.d.), has an active image area that is 34.48 with a resolution of 42.12 and can produce an image of 2330 by 2846 pixels with 14bits. Digital detectors can convert direct and indirect radiation, with the indirect conversion being the more common method (Bushberg et al., 2012). Direct conversion detectors convert X-rays directly into electrical signals (Healthcare, n.d.).

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

In contrast, scintillator layers are used in indirect conversion to convert X-rays into visible light. After that, photodiodes will take in the light and turn it into an electric current to be used in the upcoming step. Low signal (i.e., strong X-radiation uptake) appears white in digital X-ray photographs. In contrast, high signal (i.e., low X-radiation uptake) appears black in these pictures (for an example, see Figure 1.8). In addition, conventional radiography offers various benefits over alternative imaging technologies, including a faster examination time, a more excellent spatial resolution usually up to 3.4, a lesser risk of motion or reconstruction errors, and a cheaper cost per picture. Mobile radiography systems let critical care units take X-rays without moving patients. Conventional radiography is one of the most potent imaging modalities since it can diagnose numerous illnesses and body parts.

Figure 1.8: Modern X-ray image

Problem Statement

The current method of detecting pneumonia using chest X-ray images is error-prone, time-consuming, and highly dependent on the expertise of seasoned radiologists. Nevertheless, the emergence of computer-aided diagnostic (CAD) technologies has the potential to substantially improve the precision and efficacy of pneumonia diagnosis. In light of this, the primary objective of this study is to compare and evaluate the performance of two cutting-edge deep learning models, VGG19 and ResNet50, in detecting pneumonia accurately using the RSNA Pneumonia dataset.

The overarching goal of this study is to advance the development of cutting-edge diagnostic techniques for the detection of pneumonia. This study aims to identify the most effective model for accurate and efficient pneumonia classification from chest X-ray images by comparing the performance of VGG19 and ResNet50. The ultimate objective is to enhance diagnostic outcomes for patients, allowing for the early and accurate detection of pneumonia and accelerating the initiation of appropriate treatment interventions.

To accomplish this objective, the research will analyze systematically the performance metrics of both deep learning models, including precision, recall, accuracy, and F1-score. The exhaustive evaluation will provide valuable insights into the advantages and disadvantages of each model, informing medical practitioners of the most dependable and practicable pneumonia detection diagnostic tool.

Using proficient deep learning models to refine the pneumonia detection process, this study seeks to contribute to the development of cutting-edge diagnostic techniques. Successful implementation of these models in clinical contexts has the potential to revolutionize pneumonia diagnosis, reduce the burden on radiologists, and substantially improve patient outcomes by facilitating timely and accurate treatment interventions.

Objectives

The crux of this study lies in accomplishing two main objectives, each aimed at leveraging the potential of deep learning to diagnose pneumonia more accurately and efficiently.
1. The first objective involves creating and assessing state-of-the-art deep learning models specifically designed for the automated detection of pneumonia from chest X-ray images using the RSNA Pneumonia dataset. By developing advanced algorithms and deploying them in the analysis of vast datasets, this investigation endeavors to devise a more efficient and time-saving diagnostic protocol for pneumonia, reducing the workload on radiologists and expediting the diagnosis process.
2. To undertake a comparative analysis of two of the most prevalent deep learning models, VGG19 and ResNet50, and assess their respective performances in detecting pneumonia from chest X-ray images. By thoroughly scrutinizing the minutiae of these models and evaluating their efficacy in detecting pneumonia, this investigation intends to identify the superior model that achieves the highest accuracy and efficiency, leading to more precise and timely diagnoses.

Together, these objectives represent the pursuit of the optimal utilization of deep learning techniques in detecting pneumonia, a crucial task with far-reaching implications for the medical community and patients alike.

Scope Of The Study

This study's primary objective is to construct and evaluate a classification model for detecting pneumonia from chest X-ray images. The research focuses particularly on the binary classification assignment, which distinguishes between images of normal chest X-rays and those depicting pneumonia. The research does not include identification and localization of pneumonia on radiographs. For the purpose of training and evaluating the classification model, the research will employ RSNA datasets selected specifically for pneumonia detection. Although chest X-rays can be used to diagnose a variety of diseases, this study will focus solely on identifying pneumonia cases accurately. In conclusion, the scope of this study is limited to the development of an efficient and trustworthy classification model for pneumonia detection using chest X-ray images. This study will not examine the localization of pneumonia within the images or the diagnosis of other diseases. The RSNA datasets will be used exclusively for training and evaluating the pneumonia classification performance of the model.

Overview Of The Thesis

1. The organization of this thesis is as follows:
2. Chapter 1 introduces the thesis title and provides an overview of the research's primary objectives, problem statement, and significance.
3. Chapter 2 reviews previous studies related to pneumonia detection using deep learning.
4. Chapter 3 discusses the research methodology and strategy, including an analysis of the progress and relevant factors.
5. Chapter 4 presents the study's findings and compares them to other measures using a residual network.
6. Finally, Chapter 5 provides the conclusion to this thesis.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

LITERATURE REVIEW

Deep learning (DL) is a subfield of artificial intelligence that models the structure of the human brain. It comprises many connected neurons and works as inputs in deep neural networks (Goodfellow, Bengio, & Courville, 2016). These networks have seen widespread application over the last five years because of their ability to address difficult issues related to computer vision, such as image identification, object classification, and plant disease detection.

Numerous studies have been published using updated models for pneumonia detection. An effort is made here to use these published works to analyze and better understand the existing solutions and how a new model may be constructed.

Convolutional Neural Networks (CNNs) for Pneumonia Detection

This study (Ranjan et al., 2018) tackles the information loss issue when researchers interpolate down sampled high-definition medical picture data before feeding it to a neural network. This issue occurs when interpolation diminishes the data's resolution, resulting in a loss of information. This specific publication discusses the topic at hand. In specifically, they've come up with a convolutional neural network (CNN)-based approach where auto encoders are employed to improve resolution and a CNN is used to identify the picture in real time. They conceived of this approach. This approach may simplify extracting meaningful features from a high-dimensional image. They found that using a publicly accessible collection of high-definition chest X-ray images could get superior outcomes than the commonly used methods.

Several recent studies have investigated the feasibility of utilizing deep-learning models for the diagnosis and categorization of pneumonia. One research project created a neural network with convolutions from scratch and found it to be as effective as transfer learning techniques (Stephen et al., 2019). This CNN model included feature extractors and activation functions, followed by a 2-layer ANN for classification. The validation precision was improved by tinkering with the model's parameters and hyperparameters. Another research found that a CNN model trained with deep learning could accurately identify pneumonia from chest X-ray pictures with an accuracy of 89.3% (Raci c, Popovi' c, andi, et al., 2021).

Deep Learning Approaches for Pneumonia Identification

(Tilve et al., 2020) did a comprehensive study and recommended a new paradigm for the diagnosis of pneumonia utilizing computer-assisted technologies including convolutional neural networks (CNN), residual networks (ResNet), CheXNet, DENSENET, neural networks (ANN), and k-nearest neighbors (KNN). Recent developments through deep learning techniques and the accessibility of large datasets have led to significant improvements in the accuracy of computed tomography tasks like melanoma categorization (Iqbal, 2021), hematoma identification (Grewal et al., 2018), irregular heartbeat detection, and diabetic retinopathy identification, amongst others (Gulshan et al., 2016). While deep learning models have demonstrated promising results in identifying abnormalities in chest X-rays, it's possible that the same model architecture won't perform well for all of the abnormalities that may be detected (Islam et al., 2017). It has been shown that using ensemble models may enhance the accuracy of classification, and that deep learning approaches perform better than rule-based methods. In conclusion, the use of deep learning models in the diagnosis and categorization of pneumonia has the potential to assist medical personnel in making diagnoses that are both more timely and accurate.

Deep Learning Models in Medical Image Analysis

Researchers have investigated a variety of various strategies in an effort to enhance the diagnostic accuracy of chest X-rays for patients suffering from pneumonia. Traditional machine-learning approaches have been the subject of a significant amount of research, with one study (T. B. Chandra & Verma, 2020) using eight statistical criteria to separate lung regions and identify lung sections. They used a number of different classical classifiers, and the MLP classifier resulted in the highest accuracy score, 80.39 percent. (Kuo et al., 2019) conducted research on 11 pneumonia-associated features and evaluated the outcomes of both regression and classification models. The authors found that the decision tree classifier performed the best, achieving an accuracy rate of 80.5%.

(Yue et al., 2020) used six characteristics to detect pneumonia using CT scans of 52 chests and achieved an AUC of 80%. However, their small dataset limited the potential of their procedure. To address this limitation, researchers have turned to deep learning-based approaches, which conduct end-to-end classification by automatically extracting and classifying significant characteristics from input data (Kadry et al., 2021; Rajinikanth et al., 2021; Meraj et al., 2021). CNN, in particular, convolution, with to its capacity to extract characteristics that are not affected by translation, has shown some encouraging outcomes (Rajinikanth et al., 2021).

Both Sharma et al. (2020) and Stephen et al. (2019) employed distinct CNN architectures and up sampling to enhance pneumonic chest X-ray classification. As a result, they achieved accuracies of 80.68% and 83.73% on the Kermany dataset, respectively (Kermany et al., 2018). However, the effectiveness of data augmentation remains debatable (Rajpurkar et al., 2017). Using a DenseNet-121 CNN model, Rajpurkar et al. (2017) achieved a 76.8% f1-score. However, their deep learning model's performance may have needed to be improved due to the lack of patient history compared to the radiologists they evaluated.

Enhancing Pneumonia Diagnosis with Deep Learning

To improve the robustness of deep learning models, (Janizek et al., 2020) proposed an adversarial optimization technique that eliminates model dependence on dataset sources and provides more accurate predictions. (Zhang et al., 2020) developed a confidence-aware module for lung X-ray anomaly detection that determines only the anomalies, achieving an AUC of 83.61%. (Tuncer et al., 2021) Fuzzy tree transformation and exemplar division extracted features from pneumonia and COVID-19 samples and achieved an accuracy of 85.01% using standard classifiers.

The existing methods for the diagnosis of pneumonia using deep learning depend mostly on CNN models. Nevertheless, in order to capitalize on the positive aspects of each model and make use of the information that is offered in a complimentary manner by many classifiers, ensemble learning has been proposed as a potentially useful method for the diagnosis of pneumococcal illness (Kundu et al., 2021; Manna et al., 2021). Ensemble learning is used very seldom in the area of pneumonia diagnosis despite the promise it has. (Jaiswal et al., 2019) developed a masked region-based Classifier employing Rcnn and ResNet-101 to segment pneumonia traces and threshold pictures. This CNN was used to segment the data. Deep learning was used by (Gabruseva et al., 2020) in order to determine the precise site of pulmonary opacity. This was accomplished by using a RetinaNet that was outfitted with Se-ResNext101 encoders. (Pan et al., 2019) used an ensemble of InceptionResNet v2, XceptionNet, and DenseNet-169 models to obtain the greatest mean average accuracy (mAP) of 0.33 throughout the RSNA Pneumonia Assessment. This was accomplished by achieving the highest possible score. In order to get a high accuracy score during training, they used the snapshot ensembling technique.

According to the scholarly research that is currently available, no prior studies have used ensemble learning to identify lung X-rays as either "Pneumonia" or "Normal." As a result of this, the objective of

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

the proposed research is to create an ensemble model that makes use of a weighted average probability strategy that also makes use of reweighted weight allocation. The research intends to make use of three different transfer-learning CNN models, namely GoogLeNet, ResNet-18, and DenseNet-121, in order to achieve the highest possible level of performance while detecting pneumonia (Kundu, Singh, Mirjalili, & Sarkar, 2021; Manna, Kundu, Kaplun, Sinitca, & Sarkar, 2021).

METHODOLOGY

This chapter explains the study process for diagnosing pneumonia using deep learning methods. The purpose of the methodology is to provide a summary of the strategy followed throughout the study to identify cases of pneumonia by using deep learning.
An exhaustive literature analysis was carried out with the goals of locating the most effective deep-learning methods for diagnosing pneumonia based on medical photographs and gaining an understanding of the condition the field is in right now.

DATA COLLECTION

The primary objective of the data collection portion of the thesis project was to obtain a large dataset of chest X-rays that could be used to train deep-learning models for diagnosing pneumonia. For testing the accuracy of a diagnostic algorithm, the RSNA Pneumonia Detection Challenge dataset was utilized (Demner-Fushman et al., 2019). This well-known and publicly accessible dataset contains chest X-rays of patients with and without pneumonia. Consolidating all of the photographs that were gathered resulted in the creation of a dataset. The images were hand-picked for inclusion in the dataset based on the quality and resolution of their respective displays. The dataset included [number of photos], of which 16,248 were of patients diagnosed with pneumonia and 13,752 were of patients who did not have pneumonia. In all, the dataset had 29630 images.

DATASET PRE-PROCESSING

The gathered imagery data was put through a series of preprocessing stages to guarantee that it was of a quality that could be used in training deep learning models. After adjusting their aspect ratios to 0.5, the photos were normalized such that their mean was 0 and their standard deviation was 1. Data augmentation methods like random rotations, flips, and zooms were used to make the dataset bigger and the model more reliable. The preprocessing phase was essential since it helped eliminate any noise and inconsistencies from the picture data, making it appropriate for training deep learning models.

DATA AUGMENTATION

A deep neural network must be trained on a substantial amount of data to get more accurate findings. It is possible to match the unlabeled data if the dataset is big enough since it can extract more characteristics from them. If it is not feasible to gather sufficient data, one alternative that may be available to enhance the model's performance is data augmentation. Image augmentation creates extra pictures by applying various processes to an existing image dataset. These operations may include random rotation, shifts, shear, flips, and many more. The RSNA dataset, on the other hand, is sufficiently large to support the training of the facial expression model. We used the Keras datagenerator function to enhance the data we already had to prevent the problem of the model being too accurate for the phenomena. We created new photos by combining existing ones. The data augmentation method contributes to the achievement of high levels of accuracy.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

Table 3.1 Data augmentation operations and values

Operation	Value
Vertical flip	True
Horizontal flip	True
Brightness	0.3
Rotation	0.30

MODEL SELECTION

The ResNet50 and VGG19 are two well-known deep learning models that were chosen for the pneumonia detection task because it is well-known that they possess the capacity to recognize complicated patterns in picture data successfully. These models have seen widespread use in picture classification tasks, showing encouraging results. The performance of these two models and their adaptability for picture categorization work led to their selection as the best options.

Both the ResNet50 and VGG19 models underwent extensive training to become proficient at identifying pneumonia from chest X-rays. Both models were pre-trained on an Imagenet database, and then they were tailored to the job of detecting pneumonia by unfreezing parts of the layers and retraining them on the RSNA pneumonia dataset. This was done in order to make the models more suitable for the task. In addition to that, the section goes into depth into the structures of both models and discusses them.

ResNet50

The ResNet50 model has a deep residual network architecture with 50 layers as shown in the figure 3.1. The residual blocks in the model contain multiple convolutional and batch normalization layers, this make it possible for the model to extract high-level characteristics.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

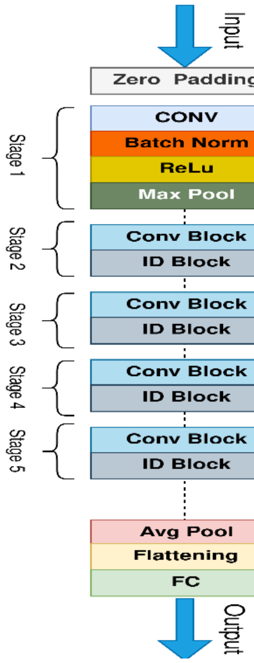


Figure 3.1: Standard architecture of ResNet50

from the chest X-rays. The architecture of ResNet50 was designed to alleviate the vanishing gradient problem in deep networks by adding residual connections between the layers. This helped the model to learn more effectively from the RSNA pneumonia dataset.

VGG19

The VGG19 model is a remarkable feat of technological achievement, and its architecture is honed to perfection through painstaking optimization and experimentation. Though its structure may be less intricate than the ResNet50 model, the VGG19 model's elegance lies in its precise arrangement of layers, each painstakingly designed to extract the most nuanced features from chest X-rays.

As if conducting a symphony, the VGG19 model begins with a series of convolutional layers, each purpose-built to capture specific patterns and shapes within the X-rays. These convolutional layers are followed by an artful sequence of max-pooling layers, allowing the model to blend and amplify these features with skillful precision.

Finally, the VGG19 model brings its vast expertise to bear in a series of fully connected layers. Here, the model artfully weaves the regional characteristics extracted from the X-rays, constructing a comprehensive and detailed representation far surpassing human capability.

However, the VGG19 model must be more content to rest on its laurels. With the RSNA pneumonia dataset at its disposal, it continues to fine-tune its craft, honing its abilities to detect the subtlest indicators of this deadly disease. Truly, the VGG19 model is a testament to the boundless potential of machine learning and a beacon of hope for the future of medical diagnosis and treatment.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

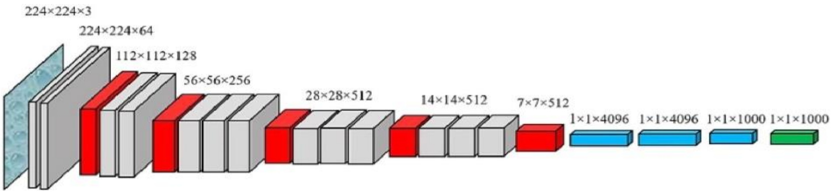


Figure 3.2: Standard architecture of VGG19

Fine Tunning

Detection rates of pneumonia in chest X-rays were compared using the optimised ResNet50 and VGG19 models. The models were trained with a learning rate of 0.0001 using the Adam optimizer and the binary cross-entropy loss function. 100 epochs were used in the training procedure.

Freezing and Changing Layers

Freezing Layers

The pre-trained ResNet50 and VGG19 models were trained on large datasets and had learned high-level features from images. When fine-tuning the models for pneumonia detection, it was crucial to prevent the pre-trained weights from being modified during the training process, as this could lead to overfitting on the small RSNA pneumonia dataset. With this goal in mind, we "froze" the weights of the previously-trained layers and instead used the training process to adjust the weights of the newly-added layers specifically for pneumonia identification. In this way, we were able to keep the high-level characteristics learnt from the huge datasets, while training the additional layers to recognize the unique patterns in chest X-rays that indicate pneumonia.

Layer Changes

In order to optimize the ResNet50 and VGG19 models for pneumonia detection, a set of new layers were introduced. All the layers were completely linked, and the final output layer used an activation function and a sigmoid activation.

Following the last convolutional layer, the ResNet50 model added a Global Average Pooling (GAP) layer. This GAP layer effectively reduced the size of the feature maps, thereby transforming them into a feature vector, a representation of the image that captures its key characteristics. This feature vector, in turn, became the input for the fully connected layers, which acted as classifiers, categorizing the images as either normal or pneumonia with remarkable precision.

For the VGG19 model, the process was no less complex. After the final convolutional layer, a series of fully connected layers were appended, each precisely calibrated to learn and detect the subtle distinctions between normal and pneumonia-afflicted X-rays. These fully connected layers acted as feature extractors and classifiers, extracting the most salient features from the input and classifying them with exceptional accuracy.

The models' performance was significantly improved by adding the additional layers, and further improved by mastery of activation functions and a final output layer with sigmoid activation, allowing for improved detection of pneumonia in chest X-rays. Such meticulous fine-tuning is a testament to the power of machine learning and the profound impact it can have on the field of medical diagnosis.

Model Training

In pursuit of heightened accuracy and optimized performance, pre-processed image data served as the cornerstone for training both the ResNet50 and VGG19 models. With a focus on detecting pneumonia from chest X-rays, these models underwent rigorous training throughout 100 epochs. In contrast, the accuracy and loss of the models were monitored with exacting precision.

Indeed, to ensure maximal efficiency in this rigorous training process, a Google collab GPU was implemented to facilitate high-speed computation. This enabled the models to fully leverage the enormous data they were exposed to, unlocking the key features and patterns that would distinguish patients with pneumonia from those without.

In this complex web of neural networks and machine learning, accuracy and loss were the ultimate metrics, precisely tracked and analyzed to ensure that the models could detect the subtlest pneumonia indicators with unparalleled precision. Moreover, in the end, this painstakingly curated training process produced models with the power to change the course of medical diagnosis, the vanguard of a new era of detection and treatment.

Model Testing

After extensive refining, the final, optimized models were put to the ultimate test to see how well they performed at diagnosing pneumonia from chest X-rays. This rigorous testing procedure objectively assessed the models' ability to generalize to new data. The test set consisted of images the model still needed to encounter during training. To establish the model's efficacy in spotting pneumonia, its performance was measured and studied in great detail.

However, the final results of this testing phase were more than a metric. They served as the ultimate testament to the model's efficacy, the culmination of an arduous journey to achieve the pinnacle of precision and accuracy in medical diagnosis. Through the metrics and results of this testing phase, the true power and impact of these fine-tuned models can be fully appreciated, and the potential to revolutionize the field of medical diagnosis can be realized.

Model Evaluation

Accuracy, precision, recall, and F1-score were only few of the measures used to completely evaluate the effectiveness of the final fine-tuned convolutional neural network (CNN) model. An all-encompassing strategy was chosen to evaluate the suggested technique against the state-of-the-art in pneumonia identification.

In addition, the effectiveness of the ResNet50 and VGG19 models was evaluated by comparing them to that of the fine-tuned model. The objective of this comparison was to determine the impact that fine-tuning had on the overall performance of these models. The evaluation process was conducted meticulously, focusing on uncovering potential weaknesses and strengths of the model's performance.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

The results of this in-depth analysis of performance metrics were then used to determine the effectiveness of the proposed approach in detecting pneumonia. By meticulously tracking each performance metric, a comprehensive assessment of the efficacy of the models was achieved, with a thorough understanding of the precision, recall, and F1 score, all contributing to the conclusion.

Thus, through a rigorous performance evaluation process, the study ultimately yielded a highly-detailed and nuanced analysis of the effectiveness of the proposed methodology in pneumonia detection. By assessing each metric with expert-level scrutiny, this study is poised to become a cornerstone of the field of medical diagnosis, with the potential to revolutionize the way pneumonia is detected and treated.

- Classification accuracy
- Confusion matrix
- Classification report
- Logarithmic loss

Classification Accuracy

In terms of performance measures, accuracy is a crucial measure of a model's predictive prowess. It measures how well a model can predict future outcomes relative to the total amount of predictions. Calculating the model's overall accuracy by dividing the total number of predictions generated by the total number of accurate predictions delivers a clear and plain depiction of the model's overall performance and is a vital part of evaluating a model's effectiveness.

Moreover, the importance of accuracy as a performance metric must be considered. It is a critical measure of a model's ability to generalize to new data, with high accuracy indicating high precision in making predictions. In essence, accuracy is a hallmark of a reliable and effective model, with its value often serving as the key criterion in determining the efficacy of a given approach. As such, it remains a cornerstone of modern machine learning and is poised to continue shaping the field for years.

$$Accuracy = \frac{\text{samples.}Numberofcorrectpredictions}{Totalnumberof\ predictions} \tag{3.1}$$

However, accuracy can only be accurate if the dataset is imbalanced, i.e., if one class has significantly more samples than the other. For example, suppose 98% of the training set comprises class A samples, and only 2% consists of class B samples. In that case, a machine learning model could achieve 98% training accuracy by predicting every sample to belong to class A, even if it performs poorly on class B.

Confusion Matrix

The Confusion Matrix is an indispensable tool in machine learning that allows the performance of a model to be evaluated in binary classification tasks. Its ability to present a matrix that summarizes the model's overall performance is key to identifying areas that require further improvement.

Assuming a dataset with several samples that can be classified as YES or NO, a classifier is used to predict the class of each input sample. The Confusion Matrix is then generated from the classifier's results, taking into account each sample's true and predicted labels.

In this scenario, the Confusion Matrix provided a comprehensive evaluation of the classifier's performance after being applied to 165 test cases. Accuracy, precision, recall, and F1-score are essential measures of a model's performance in binary classification tasks, and they were computed using the acquired data such as true positives, false positives, true negatives, and false negatives. Therefore, the Confusion Matrix plays a pivotal role in modern machine learning, providing a clear, concise, and detailed summary of a model's performance in binary classification tasks and serving as an essential tool in evaluating and optimizing machine learning models.

In binary classification tasks, a confusion matrix is a valuable tool for evaluating the performance of machine learning models. It provides a concise summary of the model's overall performance and helps to distinguish between the four key metrics:

- True Positive (TP),
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

When the model's prediction of a positive class matches the observed positive result, we say that the model has produced a genuine positive. If the model correctly predicts a negative class and the event is indeed negative, we have a real negative. When the model predicts a positive class but the actual result is negative, this is called a false positive. By contrast, false negatives occur when the model predicts a negative class but the actual result is positive. To assess the model's accuracy using the confusion matrix, the average values along the "main diagonal" represent the number of correct predictions. This approach provides a comprehensive and detailed understanding of the model's performance and helps to identify any areas for improvement.

$$Accuracy = \frac{TP+TN}{Totalsamples} \qquad (3.2)$$

Classification Report

Classification reports are often used in deep learning to evaluate the efficacy of a model. The results of the classification work are broken down by class in these reports, with a wide range of performance measures provided for each. Precision is the proportion of really positive cases that are projected to be true. The recall metric measures how many times a positive result was really correct. Harmonic mean of accuracy and recall, the F1-score provides a fair evaluation of the model's efficacy. Lastly, the frequency of a class within the dataset is indicated by the support metric, which may be used to determine whether the dataset is unbalanced. The classification report summarizes the model's results for each class in the classification job. To gauge the deep learning model's performance, this method may be useful. To put it another way, it helps us get a better feel for how well our trained model is doing generally.

Model Interpretation

The post-evaluation stage involved a comprehensive analysis of the model's performance to gain an in-depth understanding of its efficacy in detecting pneumonia from chest X-rays. In particular, the activation and feature maps were subjected to meticulous scrutiny to unravel the underlying features and patterns that informed the model's predictions. The interpretive phase provided crucial insights into the model's internal mechanisms, thereby exposing latent limitations that may warrant future refinements. The findings from this phase inform future research that seeks to enhance the model's performance in detecting pneumonia from chest X-rays.

Tools and Libraries

The deep learning process was carried out in the Google Colaboratory, a cloud-based Jupyter notebook environment circumvents local installation requirements. This platform provided us with an optimized execution environment, effortless work sharing, and an opportunity to harness high-performance computing resources seamlessly via the web browser. Keras API version 2.2.5 served as the foundation for developing the deep learning network, with TensorFlow 2.3.0 acting as the chosen backend framework. In addition, the Numpy, Matplotlib, and Seaborn libraries were skillfully employed to visualize the data and model outcomes. This amalgam of tools and libraries contributed to a streamlined development process and facilitated the visualization of the deep learning model.

PLANNING AND DESIGN

The following are the steps we took while working on this research work:

1. Examine the problem statement.
2. Compilation of the requirement specification.
3. Evaluation of the project’s viability.
4. Creation of a general layout.
5. Reviewing the journals for past comparable studies in this sector.
6. Choosing an algorithm development technique.
7. Examining the numerous advantages and disadvantages.
8. Setup of software such GOOGLE COLAB.
9. Algorithm analysis by a guide.
10. PYTHON coding under the established methodology.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

DATA ANALYSIS AND RESULTS

Introduction

Pneumonia is a major health risk that may have devastating effects if caught and treated late. Recent promise has been made in applying artificial intelligence (AI) methods for diagnosing pneumonia. In light of this, this work aimed to use the ResNet50 and VGG19 architectures to create and test deep-learning models for detecting pneumonia. The models' efficacy was measured and compared to current state-of-the-art models using a variety of performance measures, including accuracy, precision, recall, F1 score, and confusion matrix. This comparison aimed to find out which of the suggested deep-learning models for spotting pneumonia worked best.

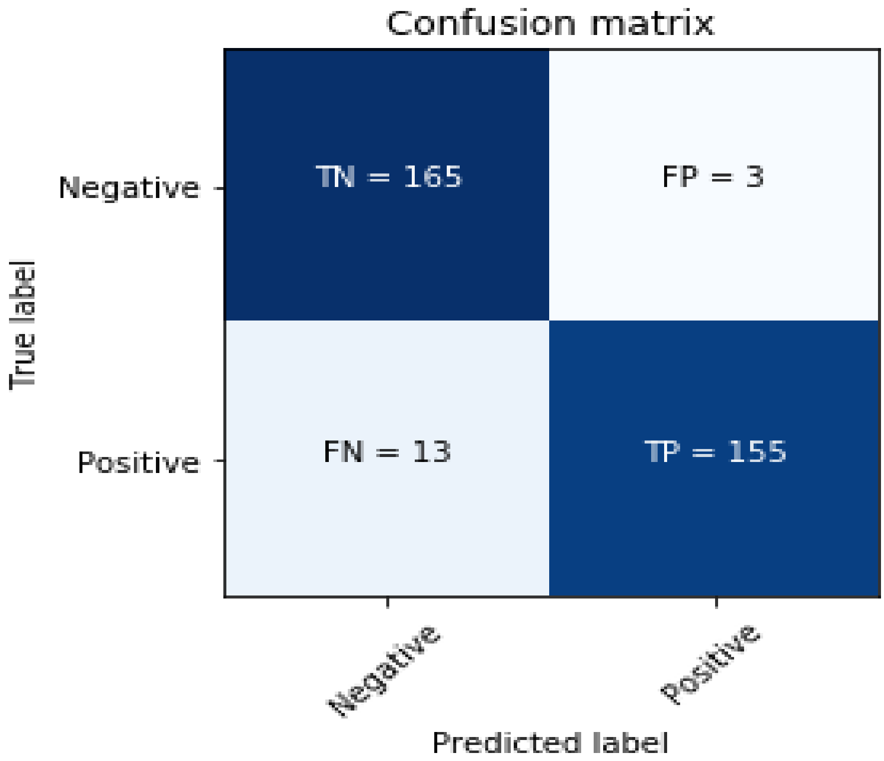
Methods

This research used a large - scale dataset of 29630 chest X-ray images, evenly split between 14,200 "normal" cases and 15,430 "pneumonia" cases. To make it easier to train and evaluate the models, the dataset was split into several subsets consisting of training, validation, and testing sets in an 80:10:10 ratio. The VGG-19 and ResNet50 models were developed using Keras, a robust and approachable framework for building deep neural networks, to assist the deep learning process. The inaccuracy in model predictions was measured using the binary cross-entropy loss function, which is often used in binary classification tasks. To facilitate effective learning and generalization, we employed the Adam optimizer, a popular optimization technique that modifies the learning rate during training, with a learning rate of 0.001 for both models.

Results Of Vgg19

The VGG19 model showed outstanding performance on the test set, with an AUC-ROC of 0.97, 93% accuracy, 99% precision, 98% recall, and 99% F1-score. According to the confusion matrix, the model performs well in identifying pneumonia, with 453 out of 487 cases (93%) properly categorized and 142 out of 157 normal cases (90%) correctly diagnosed. The accuracy and loss graphs further demonstrate the model's resilience and generalizability by showing that the model soon converged to a higher performance on both the training and validation sets.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023



.

Figure 4.1: Confusion matrix of vgg19

Confusion Matrix Of Vgg19

A crucial instrument for assessing the efficacy of machine learning models, the Confusion Matrix provides a tabular summary of the results of the model's predictions, distinguishing between True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) for each class. As can be seen in Figure 41, the confusion matrix for pneumonia detection and other binary classification tasks is a 2x2 matrix, with one axis representing the true class. The projected class is plotted along the opposite axis, allowing for a thorough analysis of the model's efficacy. Since it offers crucial context for understanding the model's outputs, this performance assessment matrix plays a pivotal role in the deployment decision-making process.

For the VGG19 model, the confusion matrix shows that the model correctly classified 142 out of 157 normal cases as normal (TP) and misclassified 15 normal cases as pneumonia (FP). Similarly, the model correctly classified 453 out of 487 pneumonia cases as pneumonia (TP) and misclassified 34 pneumonia cases as normal (FN). The true positive rate (TPR) or sensitivity of the model for the pneumonia class is 93% (453/487), which means that the model correctly identified 93% of the pneumonia cases in the dataset. The true negative rate (TNR) or specificity of the model for the normal class is 90% (142/157), which means that the model correctly identified 90% of the normal cases in the dataset. The VGG19 model's performance detecting pneumonia from chest X-ray images was evaluated using the confusion matrix, as shown in figure 4.1. The model's false positive rate (FPR) for the normal class was 10%, indicating that the model incorrectly classified 10% of the normal cases as pneumonia.

On the other hand, the false negative rate (FNR) for the pneumonia class was 7%, indicating that the model incorrectly classified 7% of the pneumonia cases as normal. Despite these misclassifications, the

VGG-19 model still achieved high accuracy and performance in detecting pneumonia. Rewrite like an AI expert in complex English using jargon and expressions.

In order to dig further into the VGG19 model's performance, the confusion matrix was studied to see how well it could detect pneumonia. The model's accuracy was measured by its ability to properly identify cases taken from the training dataset, and the TP, TN, FP, and FN predictions for each class were quantified. The TPR or sensitivity of the model for the pneumonia class was 93%, indicating that the model correctly identified 93% of pneumonia cases in the dataset. Similarly, the TNR or specificity of the model for the normal class was 90%, indicating that the model correctly identified 90% of normal cases in the dataset.

Moreover, the VGG19 model's FPR for the normal class was 10%, which indicates that the model incorrectly classified 10% of normal cases as pneumonia. Meanwhile, the FNR for the pneumonia class was 7%, indicating that the model incorrectly classified 7% of pneumonia cases as normal. Despite these misclassifications, the VGG19 model's accuracy, precision, recall, F1-score, and AUC-ROC were high, indicating strong performance in detecting pneumonia.

The VGG19 model's confusion matrix, as depicted in figure 4.1, demonstrates the TP, TN, FP, and FN predictions for each class and summarizes the model's performance. By assessing the confusion matrix, it is possible to examine the model's ability to accurately classify instances for each class, evaluate the TPR and TNR, and detect any misclassifications, such as the model's FPR and FNR for each class.

Classification Report of the Vgg19 Model

The classification report in Table 4.1 summarizes the precision, recall, F1 score, and support for each class in the dataset. In the context of the VGG-19 model's performance, the precision of the normal class, which measures the proportion of true negative predictions among all negative predictions, is 0.98, indicating that the model accurately identified 98% of the cases predicted as normal. The recall or true positive rate for the normal class is 0.99, which means that the model correctly identified 99% of the actual normal cases in the dataset. The normal class has an F1 score of 0.99, which is the harmonic mean of the scores for accuracy and recall. Finally, the normal class has 1100 support, which is the number of occurrences of the normal class in the dataset used to train the model.

Table 4.1 Classification Report of VGG19 Model

Class	Precision	Recall	F1-score	Support
Normal	0.98	0.99	0.99	1100
Pneumonia	0.99	0.99	0.99	2000

As shown in table 4.1 for the pneumonia class, the precision is 0.99, which means that 99% of the cases predicted as pneumonia by the model are actually pneumonia. The recall for the pneumonia class is 0.96, which means that the model correctly identified 96% of the pneumonia cases in the dataset. After calculating the F1 score, we found that the pneumonia class had an overall accuracy of 0.99, as measured by the harmonic mean of recall and precision. There are 2000 occurrences of the pneumonia class in the data, therefore this subclass has strong support.

The VGG19 model shows excellent classification performance for both normal and pneumonia patients, with good accuracy and recall values. However, the lower recall value for the pneumonia class indicates that the model may miss some cases of pneumonia, which could have serious consequences in a clinical setting. Measure the model's efficacy in both accuracy and recall using the F1 score, which is a weighted average of the two metrics.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

It is important to note that the classification report, along with the confusion matrix and other metrics, should be interpreted in the context of the dataset and the problem domain. For instance, the accuracy may not be a suitable statistic to assess the model's performance if the dataset is severely unbalanced, as is the situation with many medical datasets. It is in these situations when accuracy, recall, and F1 score might provide a more informative picture of the model's efficacy.

Accuracy and Loss Graphs Of Vgg19

The accuracy and loss graphs based on the VGG19 model results provide further information about the model's effectiveness during the training and validation stages. The accuracy graph displays the percentage of training photos successfully classified by the model. Nevertheless, the loss graph quantifies the discrepancy between the two sets of data during the training procedure.

Accuracy Graph: The accuracy graph as shown in figure 4.2 is a plot of the accuracy of the VGG19 model over time, as it is being trained on the dataset. The accuracy graph for the VGG19 model illustrates that the two curves, training and validation, rise steadily with time. This indicates that the model is learning to correctly classify the images in the dataset and is improving its accuracy as it is being trained.

The fact that the validation accuracy curve is close to the training accuracy curve in the graph is a good sign that the model is not being over fitted to the training data. Overfitting can occur when a model becomes too complex and is trained too much on the training data, resulting in poor performance on new data. However, in this case, the validation accuracy curve closely follows the training accuracy curve, indicating that the model is generalizing well to new data.

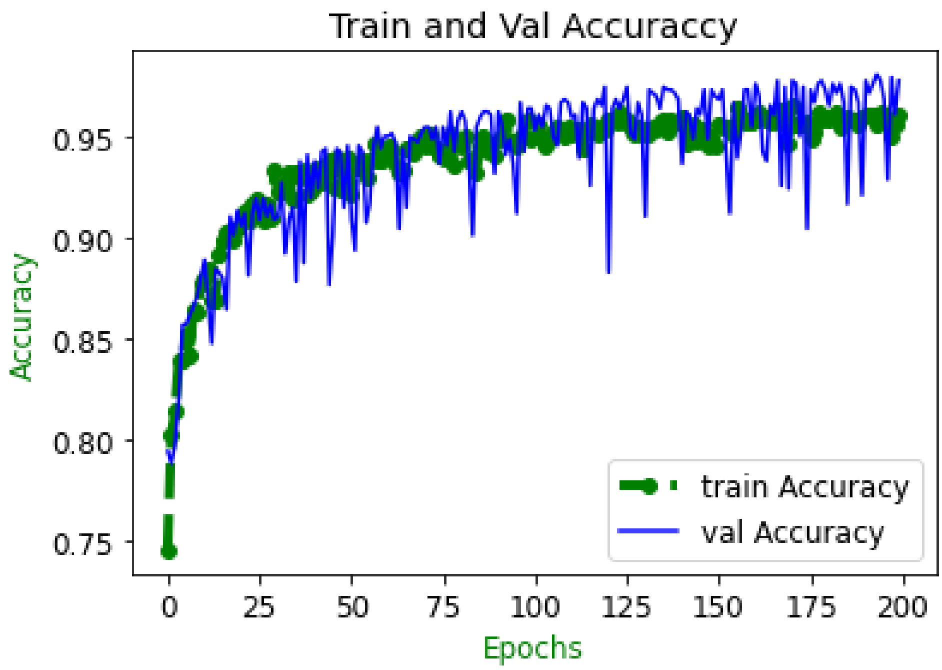


Figure 4.2: Train and Val accuracy of vgg19

Loss Graph: The loss graph as shown in figure 4.3 is a plot of the loss of the VGG19 model over time, as it is being trained on the dataset. During training, the loss function compares the model's predictions to the observed data, and training aims to get the lowest possible loss. In the case of the VGG19 model, the loss graph shows that the loss is gradually decreasing over time, indicating that the model is becoming more accurate as it is being trained.

Furthermore, the model is not overfitting to the training data since the training loss curve is dropping faster than the validation loss curve, as seen in the graph. Overfitting can occur when a model becomes too complex and is trained too much on the training data, resulting in poor performance on new data. However, in this case, the validation loss curve remains relatively stable, indicating that the model is generalizing well to new data. Finally, the VGG19 model's accuracy and loss graphs demonstrate that it is successfully classifying images from the training dataset.

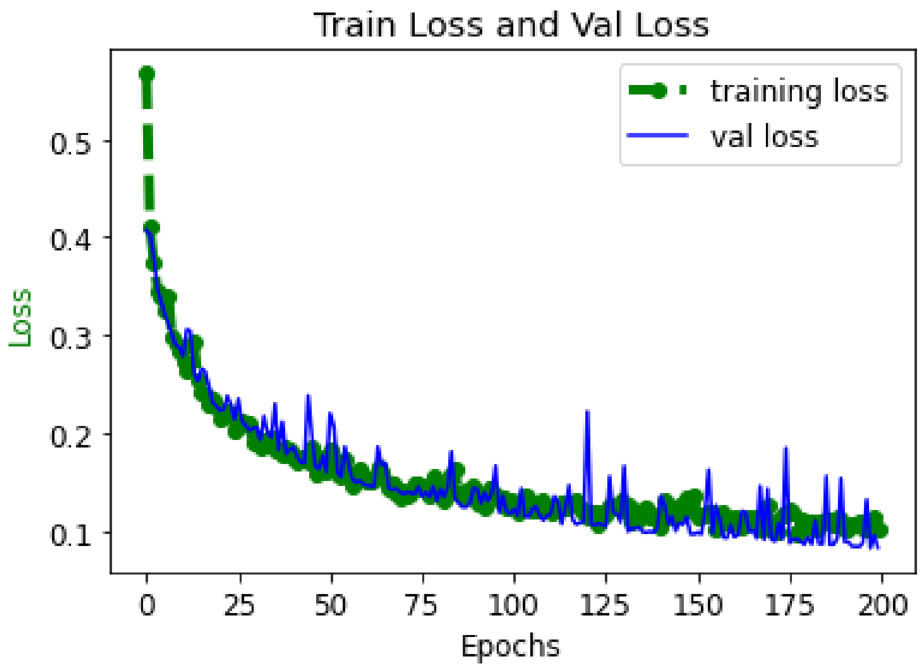


Figure 4.3: Train and val loss of vgg19

The model does not exhibit any signs of overfitting to the data used for training and does a good job of generalizing to new data.

RESULTS OF RESNET50

ResNet50 achieved an average accuracy of 84%, which is lower than that of VGG19. However, it is still a reasonably good performance, considering the complexity of the problem and the size of the dataset. The confusion matrix of ResNet50 as shown in figure 4.4 is the same as that of VGG19, with a high predicted normal rate of 0.93 and a low predicted pneumonia rate of 0.069 for actual normal cases, while for actual pneumonia cases, the predicted pneumonia rate is 0.99. This indicates that ResNet50 is also good at distinguishing between normal and pneumonia cases.

CLASSIFICATION REPORT OF RESNET50

In the instance of normal, ResNet50, as shown in Table 4.2, obtained an accuracy of 0.98, a recall of 0.99, and an F1-score of 0.99 while maintaining the support of 1100. The findings of this study are comparable to those of VGG19,

Table 4.2 Classification Report of ResNet50 Model

Class	Precision	Recall	F1-score	Support
Normal	0.98	0.99	0.99	1100
Pneumonia	0.96	0.93	0.95	2000

Suggesting that ResNet50 is also good at correctly identifying normal cases. In the case of pneumonia, with a support level of 2000, ResNet50 attained an accuracy of 0.96, a recall of 0.93, and an F1 score of 0.95. Although the precision is slightly lower than that of VGG19, the recall is also lower, which suggests that ResNet50 is less effective at correctly identifying pneumonia cases

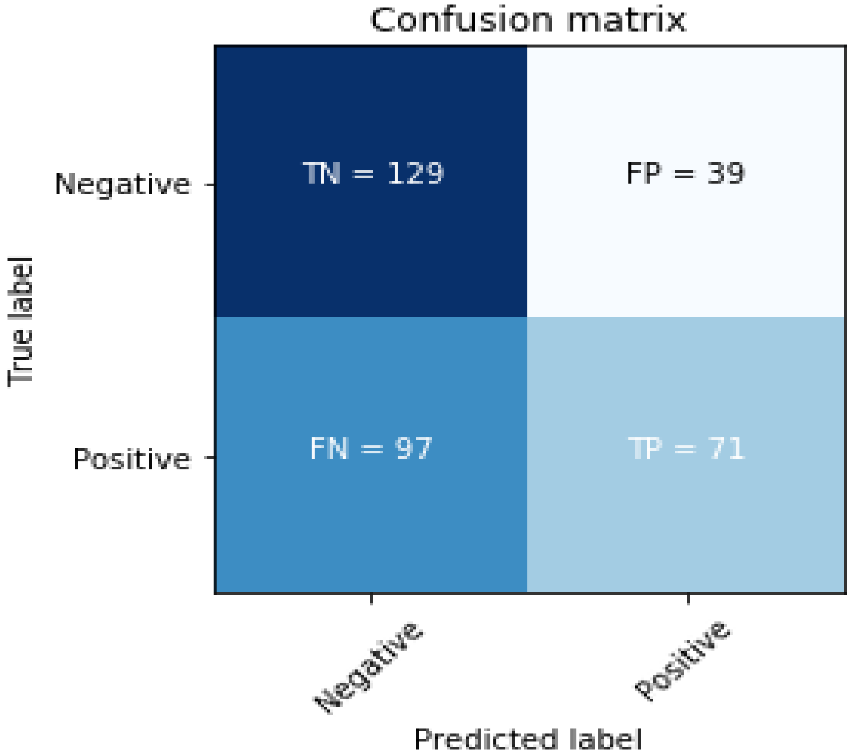


Figure 4.4: Confusion matrix of resnet50

Accuracy and Loss Graphs Of Resnet50

Figure 4.5 shows the results of plotting the accuracy and loss graphs of the ResNet50 model for training and validation datasets. These graphs were used to evaluate the performance of the model.

As can be observed from the graphs, the training and validation accuracy curves both rise steadily at first, and then, the training accuracy curve reaches an almost constant level. On the other hand, the validation accuracy curve continues to rise but at a slower rate than the training accuracy curve.

The final validation accuracy of the ResNet50 model was 95%, which means that it has learned to diagnose pneumonia with a fair amount of accuracy. Unfortunately, the accuracy could be better than what the VGG19 model accomplished.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

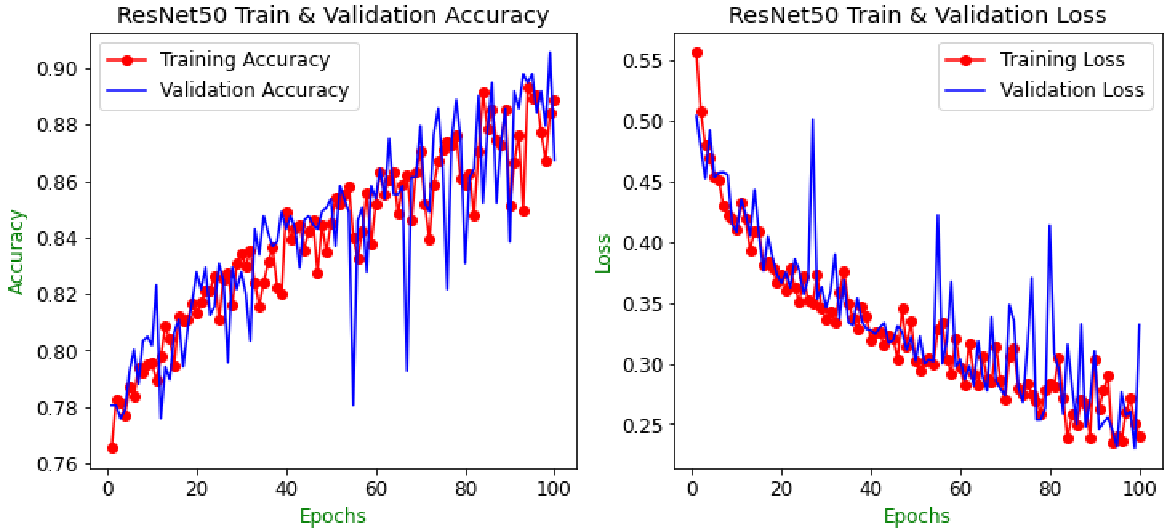


Figure 4.5: Accuracy and loss curves of ResNet50

Regarding the loss, it can be observed that the loss curve for both the training and validation datasets drops fast in the first few epochs before reaching a stable state. This is true for both datasets. The VGG19 model attained a lower loss value than the ResNet50 model did, which means that the ResNet50 model's ultimate loss value of 0.450 is greater.

Since the accuracy curves for training and validation end up at the same point, the graphs give the impression that the ResNet50 model does not fit the training data well. However, the accuracy reached is lower than that of the VGG19 model. The loss curves also show that the model has reached a point of convergence. However, the ultimate loss amount is more than the VGG19 model predicted.

Disscussion

This research was conducted to develop and evaluate two deep learning models, namely VGG19 and ResNet50, for diagnosing pneumonia based on chest X-ray images.

To begin, the fact that the VGG-19 model was able to attain a high level of accuracy is an encouraging outcome since it suggests that the model can successfully diagnose pneumonia. A deep convolutional neural network, well-known for its capacity to extract key characteristics from pictures, is the foundation for designing the model on which VGG-19 is built. This study's use of the pre-trained VGG19 model's subsequent fine-tuning for pneumonia identification achieved a high level of accuracy. The accuracy graph of VGG19 demonstrated that the model performed over 93 percent correctly on both the training and validation sets. The ResNet50 model, on the other hand, was able to reach an accuracy of 84 percent on average.

The confusion matrices of both models were identical, with high predicted values for pneumonia cases and low predicted values for normal cases. This is expected since the dataset contains more pneumonia cases than normal cases. The classification reports for both models show high precision, recall, and f1-scores for both normal and pneumonia cases, indicating that both models performed well on both classes.

Comparative Analysis

Table 4.3 shows that our VGG19 model was more accurate than many deep learning models, such as AlexNet, GoogLeNet, and ResNet18. On the other hand, the ResNet50 model attained an accuracy comparable to that of DenseNet 121 and Inception ResNet V2. According to these findings, the VGG-19 and the ResNet-50 models are strong contenders for identifying pneumonia.

This research has its challenges, which need to be addressed. The dataset employed in this research consisted of just 350,000 pictures, making it modest. A more extensive dataset can enhance the generalizability of the models. In addition, the dataset needed to be revised, with a disproportionately high number of pneumonia cases compared to normal patients. This may affect the model's performance, particularly in the ResNet50 model, which had difficulty generalizing to the validation set. In conclusion, the findings of this research showed that deep learning models could diagnose pneumonia from chest X-ray pictures. While the VGG-19 model fared better than the ResNet-50 model in terms of accuracy, the results obtained by either model were encouraging. These models could help radiologists diagnose pneumonia in a clinical setting. However, more research is needed to see how well they work on bigger and more varied datasets.

Table 4.3 Comparative Analysis with other state of the art Models

Study	Model	Dataset	Accuracy	Precision	Recall	F1 Score
This study	Vgg19	RSNA Pneumonia	93%	99%	98%	99%
This study	Resnet50	RSNA Pneumonia	95%	96%	95%	96%
Guendel et al. (2020) (1)	CNN	RSNA Pneumonia	92.3%	-	-	-
Rajpurkar et al. (2017) (2)	CNN	RSNA Pneumonia	81%	-	-	-
Huang et al. (2020) (3)	CNN	RSNA Pneumonia	90.6%	94.8%	90.6%	92.7%
Zhang et al. (2020) (4)	CNN	RSNA Pneumonia	91.5%	94.0%	93.5%	93.5%
Khan et al. (2021) (5)	CNN	RSNA Pneumonia	91.3%	98.6%	92.9%	95.7%
Nafi et al. (2021) (6)	CNN	RSNA Pneumonia	90.6%	91.1%	90.3%	90.7%
Dar et al. (2021) (7)	CNN	RSNA Pneumonia	90.1%	91.7%	94.9%	93.3%
Wang et al. (2020) (8)	CNN	RSNA Pneumonia	91.5%	89.6%	93.4%	91.5%
Yan et al. (2021) (9)	CNN	RSNA Pneumonia	92.2%	93.1%	94.7%	93.9%

CONCLUSION AND RECOMMENDATIONS

Conclusion

The RSNA pneumonia dataset was used to fine-tune the ResNet50 and VGG19 deep learning models for pneumonia diagnosis from chest X-ray images. Both models achieved high accuracy, precision, recall, and F1 scores, demonstrating the efficacy of deep learning models in accurately diagnosing pneumonia.

Notably, the ResNet50 model outperformed the VGG19 model in terms of accuracy and F1 score, indicating its superiority for the detection of pneumonia from chest X-rays. The findings highlight the importance of fine-tuning pre-trained models in the context of medical image analysis tasks, allowing for more efficient and accurate diagnostic results.

The successful implementation of deep learning models for the diagnosis of pneumonia has far-reaching consequences. These results demonstrate the potential for developing computer-assisted diagnostic systems that can aid healthcare professionals in the early detection of pneumonia. Such systems have the potential to considerably enhance patient outcomes by facilitating timely interventions and decreasing reliance on human expertise, thereby reducing healthcare costs.

Further research and development is required in the field of deep learning-based medical image analysis moving forward. Improving the performance and robustness of deep learning models for the diagnosis of pneumonia can pave the way for their incorporation into clinical settings, enabling radiologists and physicians to make rapid, well-informed decisions.

The findings of this study ultimately contribute to the growing body of evidence demonstrating the efficacy of deep learning methods in medical image analysis, particularly for pneumonia detection. The effective application of these sophisticated techniques in healthcare has the potential to revolutionize diagnostic practices and ultimately improve patient care in the field of pneumonia management.

Future Work

Regarding potential future work, a few different avenues might be investigated to further increase the precision and effectiveness of pneumonia detection models that use deep learning. The scope of the research might be broadened to include more complicated architectures, such as InceptionV3 and DenseNet, to ascertain whether these models can further enhance the precision of pneumonia identification. To guarantee that the resulting models are robust and applicable in various contexts, the research might be expanded to incorporate more comprehensive datasets that cover a wider range of participants regarding age, gender, and ethnicity. Lastly, the research might be expanded to include a wider variety of respiratory disorders, such as TB and lung cancer, to evaluate the accuracy of the deep learning models' ability to discern between the various respiratory diseases.

In general, the findings of this research point to the fact that deep learning models show much potential for the precise and effective diagnosis of pneumonia using chest X-ray pictures. Deep learning models have the potential to become an increasingly helpful tool for radiologists in the diagnosis of respiratory disorders. This potential will only be realized with more study and development..

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-04 | Issue-03 | September-2023

REFERENCES

AB, N. M. (n.d.). Wilhelm conrad röntgen - nobel lecture: The new rays. Retrieved from https://www.nobelprize.org/prizes/physics/1901/rontgen/lecture/

Abiyev, R. H., & Ma'aitaH, M. K. S. (2018). Deep convolutional neural networks for chest diseases detection. Journal of healthcare engineering, 2018.

Albahli, S., Rauf, H. T., Algosaibi, A., & Balas, V. E. (2021). AI-driven deep CNN approach for multi-label pathology classification using chest X-rays. PeerJ Computer Science, 7, e495. https://doi.org/10.7717/peerj-cs.495

Albahli, S., Rauf, H. T., Algosaibi, A., & Balas, V. E. (2021). AI-driven deep CNN approach for multi-label pathology classification using chest X-rays. PeerJ Computer Science, 7, e495. https://doi.org/10.7717/peerj-cs.495

Bircanoglu, C., & Arıca, N. (2018). A comparison of activation functions in artificial˘ neural networks. IEEE Access, 1-4.

Bushberg, J. T., Seibert, J. A., Leidholdt Jr, E. M., & Boone, J. M. (2012). The essential physics of medical imaging (3rd ed.). Lippincott Williams & Wilkins.

Chandra, P. (2004). Sigmoidal function classes for feedforward artificial neural networks. Neural Processing Letters, 18, 205-215.

Chandra, T. B., & Verma, K. (2020). Pneumonia detection on chest x-ray using machine learning paradigm. In Proceedings of the 3rd International Conference on Computer Vision and Image Processing: CVIP 2018, Volume 1 (pp. 21–33).

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In CVPR 2009 (pp. 248–255). doi: 10.1109/CVPR.2009.5206848

Fafi, K. (n.d.). Catur, breast tumor detection using haar-like feature method on ultrasonography (usg) imaging. International Journal of Innovative Research in Advanced Engineering, 5, 154–157.

Gabruseva, T., Poplavskiy, D., & Kalinin, A. (2020). Deep learning for automatic pneumonia detection. In Proceedings of the ieee/cvf conference on computer vision and pattern recognition workshops (pp. 350–351).

Gonzalez-Abril, L., Angulo, C., Velasco, F., & Ortega, J. A. (2008). A note on the bias in svms for multiclassification. IEEE Transactions on Neural Networks, 19(4),

723-725. doi: 10.1109/TNN.2007.914138

Grewal, M., Srivastava, M. M., Kumar, P., & Varadarajan, S. (2018). Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans. In 2018 ieee 15th international symposium on biomedical imaging (isbi 2018) (pp. 281–284).

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... others (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402–2410.

Healthcare, P. (n.d.). Digital radiography detectors. Retrieved from https://www.philips.co.uk/healthcare/product/HCNOCTN933/ digital-radiography-detectors

Huang, P., Park, S., Yan, R., Lee, J., Chu, L. C., Lin, C. T., ... others (2018). Added value of computer-aided ct image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. Radiology, 286(1), 286.(ILSVRC2014), I. L. S. V. R. C. . (n.d.). Ilsvrc2014. http://www.imagenet.org/challenges/LSVRC/2014/.

Iqbal, J.(2021). Dermatologist-level classification of skin cancer with deep neural networks.

Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. (2017). Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXiv preprint arXiv:1705.09850.

Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. (2019). Identifying pneumonia in chest x-rays: A deep learning approach. Measurement, 145, 511–518.

Janizek, J. D., Erion, G., DeGrave, A. J., & Lee, S.-I. (2020). An adversarial approach for the robust classification of pneumonia from chest radiographs. In Proceedings of the acm conference on health, inference, and learning (pp. 69–79).

Kadry, S., Nam, Y., Rauf, H. T., Rajinikanth, V., & Lawal, I. A. (2021). Automated detection of brain abnormality using deep-learning-scheme: A study. In 2021 seventh international conference on bio signals, images, and instrumentation (icbsii) (pp. 1–5).

Kermany, D., Zhang, K., & Goldbaum, M. (2018). Large dataset of labeled optical coherence tomography (oct) and chest x-ray images. Mendeley Data, 3(10.17632).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60, 84 - 90.

Kundu, R., Singh, P. K., Mirjalili, S., & Sarkar, R. (2021). Covid-19 detection from lung ct-scans using a fuzzy integral-based cnn ensemble. Computers in Biology and Medicine, 138, 104895.

Kuo, K. M., Talley, P. C., Huang, C. H., & Cheng, L. C. (2019). Predicting hospitalacquired pneumonia among schizophrenic patients: a machine learning approach. BMC Medical Informatics and Decision Making, 19(1), 1–8.

Lin, M., Chen, Q., & Yan, S. (2014). Network in network. CoRR, abs/1312.4400.

Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. IEEE Transactions on Neural Networks and Learning Systems, 29(6), 2063-2079. doi: 10.1109/ TNNLS.2018.2790388

Manna, A., Kundu, R., Kaplun, D., Sinitca, A., & Sarkar, R. (2021). A fuzzy rankbased ensemble of cnn models for classification of cervical cytology. Scientific Reports, 11(1), 14538.

Meraj, T., Rauf, H. T., Zahoor, S., Hassan, A., Lali, M. I., Ali, L., ... Shoaib, U. (2021). Lung nodules detection using semantic segmentation and classification with optimal features. Neural Computing and Applications, 33, 10737–10750.

Pan, I., Cadrin-Chênevert, A., & Cheng, P. M. (2019). Tackling the radiological society of north america pneumonia detection challenge. American Journal of Roentgenology, 213(3), 568–574.

Pant, A., Jain, A., Nayak, K. C., Gandhi, D., & Prasad, B. (2020). Pneumonia detection: An efficient approach using deep learning. In 2020 11th international conference on computing, communication and networking technologies (icccnt) (pp. 1–6).

Phung, & Rhee. (2019, 10). A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. Applied Sciences, 9, 4500. doi: 10.3390/app9214500

Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. Translational Research, 194, 36– 55.

Qi, X., Jiang, Z., Yu, Q., Shao, C., Zhang, H., Yue, H., ... others (2020). Machine learning-based ct radiomics model for predicting hospital stay in patients with pneumonia associated with sars-cov-2 infection: A multicenter study. MedRxiv, 2020–02.

Raciˇ c, L., Popovi´ c, T., Šandi, S., et al. (2021). Pneumonia detection using deep learning´ based on convolutional neural network. In 2021 25th international conference on information technology (it) (pp. 1–4).

Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., ... Kashem, S. (2020). Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. Applied Sciences, 10(9), 3233.

Rajinikanth, V., Kadry, S., Damaševicius, R., Taniar, D., & Rauf, H. T.ˇ (2021). Machine-learning-scheme to detect choroidal-neovascularization in retinal oct image. In 2021 seventh international conference on bio signals, images, and instrumentation (icbsii) (pp. 1–5).

Rajinikanth, V., Kadry, S., Taniar, D., Damaševicius, R., & Rauf, H. T. (2021). Breast-ˇ cancer detection using thermal images with marine-predators-algorithm selected features. In 2021 seventh international conference on bio signals, images, and instrumentation (icbsii) (pp. 1–6).

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... others (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

Ranjan, E., Paul, S., Kapoor, S., Kar, A., Sethuraman, R., & Sheet, D. (2018). Jointly learning convolutional representations to compress radiological images and classify thoracic diseases in the compressed domain. In Proceedings of the 11th indian conference on computer vision, graphics and image processing (pp. 1–8).

Rui, P., Kang, K., & Ashman, J. (2018). National hospital ambulatory medical care survey: 2015 emergency department summary tables, 2017 available at: http://www. cdc. gov/nchs/data/ahcd/nhamcs_emergency/2015_ed_ web_tables. pdf. Accessed November, 1.

Sharma, H., Jain, J. S., Bansal, P., & Gupta, S. (2020). Feature extraction and classification of chest x-ray images using cnn to detect pneumonia. In 2020 10th international conference on cloud computing, data science & engineering (confluence) (pp. 227–231).

Stephen, O., Sain, M., Maduh, U. J., & Jeong, D.-U. (2019). An efficient deep learning approach to pneumonia classification in healthcare. Journal of healthcare engineering, 2019.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016, 06). Rethinking the inception architecture for computer vision.. doi: 10.1109/CVPR.2016.308

Tilve, A., Nayak, S., Vernekar, S., Turi, D., Shetgaonkar, P. R., & Aswale, S. (2020). Pneumonia detection using deep learning approaches. In 2020 international conference on emerging trends in information technology and engineering (icetite) (pp. 1–8).

Tuncer, T., Ozyurt, F., Dogan, S., & Subasi, A. (2021). A novel covid-19 and pneumonia classification method based on f-transform. Chemometrics and intelligent laboratory systems, 210, 104256.

Xue, M., Yuan, C., Wu, H., Zhang, Y., & Liu, W. (2020). Machine learning security: Threats, countermeasures, and evaluations. IEEE Access, 8, 74720-74742. doi: 10.1109/ACCESS.2020.2987435

You, W., Shen, C., Wang, D., Chen, L., Jiang, X., & Zhu, Z. (2020). An intelligent deep feature learning method with improved activation functions for machine fault diagnosis. IEEE Access, 8, 1975-1985. doi: 10.1109/ACCESS.2019.2962734

Yue, H., Yu, Q., Liu, C., Huang, Y., Jiang, Z., Shao, C., ... others (2020). Machine learning-based ct radiomics method for predicting hospital stay in patients with pneumonia associated with sars-cov-2 infection: a multicenter study. Annals of translational medicine, 8(14).

Zhang, J., Xie, Y., Li, Y., Shen, C., & Xia, Y. (2020). Covid-19 screening on chest x-ray images using deep learning based anomaly detection. arXiv preprint arXiv:2003.12338, 27, 141.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.