AUWALU SALEH MUBARAK

EFFICIENT COVID-19 DETECTION ON COMPUTED TOMOGRAPY IMAGES

NEU 2022

EFFICIENT COVID-19 DETECTION ON

COMPUTED TOMOGRAPHY IMAGES

A THESIS SUBMITTED TO THE INSTITUTE OF GRADUATE STUDIES OF NEAR EAST UNIVERSITY

By
AUWALU SALEH MUBARAK

In Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Electrical and Electronic Engineering

NICOSIA, 2022

EFFICIENT COVID-19 DETECTION ON COMPUTED TEMOGRAPY IMAGES

A THESIS SUBMITTED TO THE INSTITUTE OF GRADUATE STUDIES OF NEAR EAST UNIVERSITY

By
AUWALU SALEH MUBARAK

In Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Electrical and Electronic Engineering

NICOSIA, 2022

Auwalu Saleh Mubarak: EFFICIENT COVID-19 DETECTION ON COMPUTED TOMOGRAPHY IMAGES

Approval of Director of Institute of Graduate Studies

Prof. Dr. K. Hüsnü Can BAŞER

We certify this thesis is satisfactory for the award of the degree of Doctor of Philosophy in Electrical and Electronic Engineering

Examining Committee in Charge:

Assoc.Prof. Sertan Serte	Supervisor, Department of Electrical and Electronics Engineering, NEU
Prof. Fadi Al-Turjman F. Al-Turjman	Co-Supervisor, Department of Artificial Intelligence Engineering, NEU
Prof. Mehmet Ozsoz	Committee Chairman, Department of Biomedical Engineering, NEU
Prof. Ayse Gunay	Committee Chairman, Department of Biomedical Engineering, NEU
Assoc. Prof. Enver Ever	Department of Computer Engineering, METU
Assoc. Prof. D. B. Deebak	School of Computer Science and Engineering, VIT
Prof. Ahmet Denker Ahmet Denker	Department of Electrical and Electronics Engineering, USF

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Auwalu Saleh Mubarak

Signature:

Date:

ACKNOWLEDGEMENTS

My deepest gratitude goes to my supervisors Assos. Prof. Dr. Sertan Serte and Prof. Dr. Fadi Al-Turjman, for their exemplary guidance, patience and support during the course of this thesis. This thesis would not have been possible without their help and vast wisdom.

I would like to extend my appreciation and prayers to my parents Alhaji Auwalu Saleh and Halima Ridwan, for their sacrifice, love, care and moral support. Also, my special appreciation goes to my wife Zubaida Said Ameen for her love, prayers and patience.

Lastly, I thank my family members, and my friends for their constant encouragement and being there for me always without which this assignment would not be possible.

To my parents…

ABSTRACT

So far, the deadly Corona Virus (COVID-19), which was originally discovered in China in December 2019, has claimed the lives of almost four million people worldwide. This deadly disease has crippled educational, economic, and social activities throughout the world. During the 14-day incubation period, a person who is potentially infected with the disease may not display any symptoms. To stop the virus from spreading, the government and other organizations are spending a lot of money isolating persons who come into contact with someone who turns out to be positive. The benchmarked method of detecting the virus is 19 is the Reverse-Transcription Polymerase Chain Reaction (RT-PCR). The method is associated with human errors and a detection rate of 50% to 70%. In the diagnosis of patients, medical images like X-ray and Computed Tomography(CT) of the lungs are been used, though x-ray is associated with a lateral view, unlike CT where slice level information of the lungs can be achieved.

Medical imaging and Artificial Intelligence (AI) have proven to be highly effective in detecting lung diseases including lung cancer and tuberculosis. Because the number of cases is increasing every day throughout the pandemic, a fast and reliable method of detecting COVID-19, particularly CT scans, is critical. Handcrafted Features such as Local Binary Pattern(LBP) and automated Deep Learning Features of several pre-trained models were utilized to train SVM and KNN classifiers to detect three classes of CT scan images: COVID-19 positive, Common Pneumonia(CP), and Healthy persons. The models proposed in this research has attained a high classification capability based on the performance criteria considered in the study, also it can serve as an alternative to RT-PCR because it can detect the virus at an early stage. This will reduce the cost of isolation, improve detection and reduce the spread of the virus.

Keywords: COVID-19; Artificial Intelligence; Support Vector Machines; K-Nearest Neighbour, Deep Learning; Machine Learning.

ÖZET

Aslen Aralık 2019'da Çin'de keşfedilen ölümcül Corona Virüsü (COVID-19), şimdiye kadar dünya çapında yaklaşık dört milyon insanın hayatına mal oldu. Bu ölümcül hastalık, dünya çapında eğitim, ekonomi ve sosyal faaliyetleri sekteye uğrattı. 14 günlük kuluçka döneminde, hastalığa yakalanma potansiyeli olan bir kişi herhangi bir semptom göstermeyebilir. Virüsün yayılmasını durdurmak için hükümet ve diğer kuruluşlar, pozitif çıkan biriyle temasa geçen kişileri izole etmek için çok para harcıyor. Virüsü saptamanın kıyaslanmış yöntemi 19'dur, Ters Transkripsiyon Polimeraz Zincir Reaksiyonu (RT-PCR). Yöntem, insan hataları ve %50 ila %70'lik bir algılama oranı ile ilişkilidir. Hastaların tanısında akciğerlerin kesit düzeyinde bilgi elde edilebilen BT'den farklı olarak, röntgen lateral bir görünümle ilişkilendirilse de, akciğerlerin X-ray ve Bilgisayarlı Tomografi (BT) gibi tıbbi görüntüleri kullanılır.

Tıbbi görüntüleme ve Yapay Zekanın (AI), akciğer kanseri ve tüberküloz dahil olmak üzere akciğer hastalıklarını tespit etmede oldukça etkili olduğu kanıtlanmıştır. Pandemi boyunca vaka sayısı her geçen gün arttığından, COVID-19'u, özellikle de BT taramalarını tespit etmenin hızlı ve güvenilir bir yöntemi kritik öneme sahiptir. Yerel İkili Model (LBP) gibi El Yapımı Özellikler ve önceden eğitilmiş birkaç modelin otomatik Derin Öğrenme Özellikleri, SVM ve KNN sınıflandırıcılarını üç sınıf CT tarama görüntüsünü algılamak üzere eğitmek için kullanıldı: COVID-19 pozitif, Yaygın Pnömoni (CP) ve Sağlıklı kişiler. Bu araştırmada önerilen modeller, çalışmada dikkate alınan performans kriterlerine göre yüksek bir sınıflandırma kabiliyetine kavuşmuştur, ayrıca virüsü erken aşamada tespit edebildiği için RT-PCR'a alternatif olarak hizmet edebilir. Bu, izolasyon maliyetini azaltacak, algılamayı iyileştirecek ve virüsün yayılmasını azaltacaktır.

Anahtar kelimeler : COVID-19; Yapay zeka; Vektör makineleri desteklemek; K-En Yakın Komşu, Derin Öğrenme; Makine öğrenme.

TABLE OF CONTENT

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI:	Artificial Intelligence
CT:	Computer Tomography
COVID-19:	Corona Virus
SVM:	Support Vector Machines
K-NN:	K-Nearest Neighbour
LBP:	Local Binary Patterns
DL:	Deep Learning
ML:	Machine Learning

CHAPTER 1
INTRODUCTION

1.1 Introduction

In December 2019, a deadly respiratory disease Corona Virus 2 (SARS-CoV-2)/(COVID-19) which was different from SARS-CoV and MERS-CoV was reported in Wuhan, Hubei Province, China(Lu et al., 2020) the disease spread was suspected to have started from wild animals which are illegally sold in a seafood market in Wuhan, in few days the deadly disease spread to many cities in China and other countries (Zhu et al., 2020). On 30th January 2020, World Health Organization declared SARS-CoV-2 as a global pandemic (Ye et al., 2020), By 16th of march 2020, the number of active and death cases reported in the world was 168,826 and 6,503(Zhai et al., 2020), and by the 27th of August 2020 the confirmed cases reaches 24,514,320 and total death of 832,660, the rate of infection is increasing exponentially. The United States, Spain, Italy, the United Kingdom, and France are among the countries with the highest number of infected persons. The elderly and those with weakened immune systems are the most vulnerable (Tuli et al., 2020). After the current epidemic, it is predicted to be a seasonal illness (Neher et al., 2020). New cases are disclosed on a regular basis all across the world (Rustam et al., 2020). The spread of the virus has sparked fear and panic throughout the world, causing nations to implement entire or partial lockdowns (Chamola et al., 2020). Non-pharmaceutical interventions (NPIs) such as social distancing, school closure, and voluntary quarantine have all been shown to be effective in reducing illness spread (Tuli et al., 2020) (Oruc et al., 2021). These techniques, on the other hand, are detrimental to daily life, the economy, business, and education. Due to the closure of schools, colleges, and institutions, e-Learning is expected to become the norm in many nations throughout the world. However, there are other obstacles to overcome, including a lack of training for faculty members on Internet-based teaching approaches because they are used to traditional ways. Furthermore, owing to the lack of availability or inadequacy of the Internet in rural places, this strategy is extremely difficult to implement. Furthermore, e-learning has the disadvantage of requiring pupils to commit to studying. As a result, both students and instructors are harmed by this pandemic (Oruc et al., 2021) (Murphy, 2020)(Williamson et al., 2020). As a result, parents should assist their children in learning their lessons remotely to keep them busy and cognitively interested during the quarantine time (Williamson et al., 2020). Furthermore, quarantine has a significant impact on mental health, including psychological and emotional disorders(Ammar et al., 2020). The common symptoms of the disease include fever, dry cough, fatigue, lymphopenia and acute respiratory distress with severe pneumonia(D. Wang et al., 2020). A very rapid and efficient method of detecting COVID-19 is needed to curve the spread of the virus as it affects the whole world.

1.2 Background of Study

This Coronavirus is a novel virus that causes COVID-19 disease. It is also called SARS-CoV-2 (Tuli et al., 2020). The disease has first appeared in Wuhan, China, in December 2019 and has spread worldwide (Mohammed, Syamsudin, et al., 2020). A worldwide health emergency has been declared due to the COVID-19 disease epidemic (Mohammed, Hazairin, et al., 2020). Fever, cough, and breathing problems are the most prevalent respiratory signs of this condition (Mohammed, Hazairin, et al., 2020). Symptoms of acute respiratory infection appear early in the illness, and patients may progress quickly to the advanced stage, developing acute respiratory distress syndrome (ARDS) (N. Chen et al., 2020a). It is critical to classify the new COVID-19 as soon as possible to treat and control the disease(Singh et al., 2020). The laboratory test has several disadvantages, including the necessity for a long period and a significant cost. It has also been discovered that RT-PCR detection of viral RNA from sputum or a nasopharyngeal swab has a low positive rate in the early stages of COVID-19 identification (Butt et al., 2020). CT diagnosis is a quick method of diagnosis, although doctors' ability to diagnose the disease using this method is just average. Furthermore, this procedure is inefficient, especially when there are a large number of patients. As a result, new solutions to assist them in the diagnosis are required (Abbasian et al., 2020)(Nguyen et al., 2020). Early detection of coronavirus infection and following the condition of affected persons are both aided by artificial intelligence. It also aids in the facilitation of viral research by assessing accessible data. Healthcare organizations require decision-making tools to cope with this virus and to aid them in receiving appropriate recommendations in real-time to stop it from spreading.

Artificial intelligence (AI) effectively imitates human intelligence. It's also beneficial for comprehending and offering improvements to a COVID-19 vaccination (Vaishya et al., 2020). In the current circumstances, all nations are seeking for effective and low-cost solutions to assist them in dealing with this disease (Pratap et al., 2020). Computer scientists have been able to detect the disease early using AI skills by evaluating medical image data (Nguyen et al., 2020) (Alimadadi et al., 2020). ML has the benefit of being able to effectively solve complicated problems (Rustam et al., 2020). The Internet of Things (IoT) combines multiple linked devices to form an intelligent network that may be used to implement a proper health management system. It's utilized to keep track of illnesses and ensure the patient's safety. It captures the patient's medical data without requiring any human interaction (Pratap et al., 2020). Infrared thermometers are used to assess body temperature in all public settings. The risk of the virus transmitting from the infected individual to the person doing the screening technique is a disadvantage of this approach. As a result, alternate solutions with less human contacts are urgently needed (Mohammed, Hazairin, et al., 2020). Due to the widespread COVID-19 in the world, fast and early detection of COVID-19 and the low-cost test is very important, many people have been isolated because they show mild symptoms and the virus has an incubation period of 14 days(Xie & Chen, 2020). The benchmark test result for COVID-19 today is RT-PCR test for the detection of the nucleic acid forms stem from the COVID-19, the test is performed by collecting respiratory specimens such as oropharyngeal swabs, in the receiving place of specimen human error can occur(R. Liu et al., 2020). The PCR test is time-consuming and costly, patients awaiting results must be isolated which will increase cost either to the government or patient, also the chances of getting the correct result are 30% -50% (Chu et al., 2020). An alternative to PCR is medical imaging whereby Computed Tomography (CT) scan images can be applied most especially in the case of pregnant and small children(H. Liu et al., 2020; Strunk et al., 2014). Though CT is associated with high dosage and is also costly(Kroft et al., 2019) but was used for COVID-19 evaluation and profiling in (Kroft et al., 2019)(Strunk et al., 2014). (Strunk et al., 2014)(N. Chen et al., 2020b)recommended Xray than CT scans but did not consider pregnant women and children.

1.3 Research Problem and Statement

Computer-aided diagnostic (CAD) systems can help radiologists improve the accuracy rate. Researchers are currently using handcrafted or learning features that are centred on lung texture, structure, and morphological characteristics for identification. However, it is always important and difficult to select the right classifier that can optimally handle the properties of the lung spaces. The traditional image recognition methods are support vector machine (SVM), k-nearest neighbours (kNN), artificial neural networks (ANNs), decision trees (DTs) and Bayesian networks (BNs). These machine-learning methods(Fehr et al., 2015; Orrù et al., 2012) needs hand-crafted features to compute such as morphological, texture, SIFT, entropy, the density of pixels, elliptic shape, geometry, Fourier descriptors (EFDs), and off-shelf classifiers as explained in (X. Chen et al., 2021). In comparison, machine-based learning (ML) approaches are known as non-deep learning methods. There are many uses for these non-deep learning approaches, such as the use of neurodegenerative diseases, cancer diagnosis, and psychological disorders (Cruz & Wishart, 2006; Doyle et al., 2007; Oakden-Rayner et al., 2017; Orrù et al., 2012; Parmar et al., 2017). However, the major drawbacks of non-deep learning approaches are that they rely on the extraction stage of the function and this makes it challenging to find the most important feature required to produce the most successful outcome. The application of artificial intelligence (AI) will be used to solve these difficulties. AI technology in the field of medical imaging is becoming increasingly popular, particularly for the advancement of technology and the growth of deep learning. (Gao et al., 2017; Sa et al., 2021; J. Wang et al., 2017; Q. Zhang et al., 2016).

Convolutional neural networks (CNNs) have attained state-of-the-art performance in the field of medical imaging based on previous studies (Waheed et al., 2020)(Shuihua Wang, Sun, et al., 2019)(Shuihua Wang, Tang, et al., 2019). This level of reliability is achieved by training and fine-tuning the system's millions of parameters with labelled data. Because of the large number of parameters, CNN can easily overfit small datasets, so generalization efficiency is proportional to the size of the labelled data. Due to the limited number of datasets, limited datasets prove to be the most challenging problem in the medical imaging domain (Roth et al., 2015)(Cnns, 2016)(Tajbakhsh et al., 2016). Medical image acquisition is a very expensive and tedious process that requires the participation of radiologists and researchers (Cnns, 2016). Furthermore, due to the recent severity of the COVID-19 disease, sufficient data of chest CT scan images are tough to obtain, unlike in (Waheed et al., 2020) whereby the detection of COVID-19 was performed on synthetic CT scan images, we proposed an offline data augmentation were by several data augmentations employed in many studies were performed on each of the three datasets employed in the study such as random reflection, random rotation, random rescale, random translation.

Feature extraction is a critical component of a detection system's performance (Niu & Suen, 2012). CNN features are automatically trained. One of CNN's advantages in the case of transformations such as translation, scaling, and rotation is that they can be invariant. Invariance, rotation, and scale are three of CNN's most unique advantages, particularly in image recognition problems such as object detection, since they allow the network to abstract identity, enabling it to recognize the object even though the image's pixel values vary greatly. Feature extraction increases the accuracy of the models learned by extracting the features from the input data. This move in the general framework reduces data dimensionality by removing redundant data. It also increases the speed and inference of model training. Methods of extraction of features generate new features by rendering the variations and transformations of the original features. Colour, shape, texture or pixel value is the type of characteristics that can be obtained from medical images. Any diagnostic image, such as CT scan images, does not contain any colour detail. This is appreciated in this field.

.

1.4 Research Aim and Objectives

This research aims at efficiently detecting COVID-19 by analyzing the performance of different ML classifiers such as SVM and KNN on handcrafted and automatic Deep Learning features to detect the COVID-19 at an early stage and reduce the spread of the virus.

Following research objectives would facilitate the accomplishment of this aim:

- Proposed a COVID-19 detection model by extracting seven different learning model features and using SVM as a classifier.
- Proposed a model that will efficiently detect COVID-19 Using Local Binary Pattern and Deep Learning feature fusion on SVM and KNN classifiers

1.5 Scope of the Study

The The emphasis of this study is to efficiently detect COVID-19 by classifying CT scan images of three different classes namely COVID-19, CP and Healthy patients. The study will be based on the effect of the features extracted deep learning features and handcrafted features. The extracted features will be used by different classifiers to classify the CT scan images. The whole modelling is carried out in a MATLAB 2019b environment

1.6 Thesis Organization

The thesis report is organized into five chapters; and the research objectives are thoroughly elaborated in chapters 3 and 5, which can be read independently from each other. The chapters are arranged as follows:

Chapter 1: In this chapter, the background of the study is presented. The chapter also discusses the problem statement, objective and scope of the study.

Chapter 2: Reviews basic artificial intelligence, deep learning, transfer learning, medical imaging. It also provides a thorough and concise literature review of the other related studies.

Chapter 3: This chapter illustrates the development of a new COVID-19 detection model by extracting seven different learning model features and using SVM as a classifier.

Chapter 4: This chapter described another scenario of a model that will efficiently detect COVID-19 Using Local Binary Pattern and Deep Learning feature fusion on SVM and KNN classifiers.

Chapter 5: In this chapter, the conclusion of the thesis work is presented along with recommendations for future work.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

In this chapter, the literature review on the relevant studies is presented. The chapter begins by reviewing AI applications in COVID-19 era, followed by medical imaging, transfer learning, COVID-19 detection using AI, employing machine learning in prediction, segmentation of CT images and lastly, challenges of using AI in COVID-19 detection

2.2 AI Applications in COVID-19 Era

Sustainable development is the enhancement in all aspects of life and meeting the current needs without adversely affecting the needs of future generations. Epidemics are one of the obstacles that slow down sustainable development. COVID-19 epidemic affects the economic and social aspects. Therefore, many researchers seek solutions that help to combat this epidemic (Pirouz et al., 2020). In (Pirouz et al., 2020) classification of affirmed cases of COVID-19 was done. Binary classification modelling has been employed by the group method of data handling (GMDH) type of NN which is one of the AI approaches. So, the Hubei province in China has been chosen as a case study to create the proposed model, and some significant factors which are highest, lowest, and average daily temperature, the density of a city, relative humidity, and wind speed, have been considered as the input dataset, and the number of affirmed cases has been chosen as the output dataset for one month. The proposed binary classification approach performs better in predicting confirmed instances. Furthermore, regression analysis was used to determine the trend of confirmed cases as well as variations in daily weather factors (wind, humidity, and average temperature). The findings revealed that the maximum daily temperature and relative humidity had the largest impact on the confirmed cases. The relative humidity in the main case study, which averaged 77.9 %, had a beneficial influence on the approved instances. The greatest daily temperature, with an average of 15.4 °C, had a detrimental impact on the confirmed cases.

It is critical to construct reliable forecasting models in order to take proper action. Because of the scarcity of basic data, epidemiological models were evaluated to see if they could provide higher accuracy for long-term predictions. The research presented in this publication provides an alternative to susceptible-infected-resistant (SIR) models. In (Pinter et al., 2020), a mixed ML technique has been developed to forecast the COVID-19. Data from Hungary was used. A hybrid ML technique combining adaptive network-based fuzzy inference system (ANFIS) and multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) has been proposed to forecast time series of infected people and death rate. The models foresee that at the end of May, the outbreak will fall considerably. The confirmation was carried out for 9 days with superb outcomes, which affirms the model accuracy. It is anticipated that the model will retain its accuracy. This work has exhibited the capability of ML for future research.

In research by (Zheng et al., 2020) a hybrid AI model was presented to predict COVID-19 disease. It was found that conventional models deal with all infected people as they have the same infection rate. So, an enhanced model has been proposed to predict the variation of the infection rates. To create this hybrid model, the natural language processing (NLP) module, and the long short-term memory (LSTM) network was included. The experimental findings on the epidemic data of numerous typical provinces and cities in China demonstrate that people with COVID-19 have a greater infection rate within the third to eight days after they were infected. Compared to conventional epidemic models, the proposed model can greatly decrease the errors of the prediction findings.

In (Ghoshal & Tucker, 2020), the Bayesian DL classifier was trained via a transfer learning approach on COVID-19 X-Ray images. It has been used for estimating model uncertainty. It was found that there is a significant correlation between model uncertainty and accuracy of prediction. The estimated uncertainty in DL yields a more trustworthy forecast, which can notify radiologists of false predictions. So, this will increase their trust in the ability of DL to help in disease detection.

In (Iwendi et al., 2020), A fine-tuned Random Forest model was boosted using the AdaBoost approach. This algorithm predicts the severity of the illness and the anticipated outcome—

recovery or death—based on COVID-19 patients' geographic, travel, health, and demographic data. On the used dataset, an accuracy of 94% and an F1 score of 0.86 were obtained. The data analysis demonstrates a significant correlation between patient gender and mortality, as well as the fact that the majority of patients are between the ages of 20 and 70. Male patients were shown to have a greater death rate than female ones. It was also demonstrated that the Boosted Random Forest method can predict accurately even on unbalanced datasets.

Medical imaging (Zhou et al., 2021) Non-invasive or intrusive procedures are used to make visual representations or photographs of external or inner tissues of the human body or a segment of the human body using physical phenomena such as light, electromagnetic radiation, radioactivity, nuclear magnetic resonance, and sound. X-ray radiography, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and digital pathology are the most often used imaging modalities in clinical practice. Imaging data accounts for approximately 90% of all healthcare data1, making it one of the most useful sources of evidence for clinical analysis and medical intervention.

The practicality and nature of deep learning solutions are influenced by numerous features in medical imaging. It's worth noting that these traits aren't exclusive to medical imaging. In terms of the first feature described below, satellite imaging, for example, is similar to medical imaging. Medical images have high pixel resolution and include a variety of modalities. Many imaging modalities are currently in use, and new ones, such as spectral CT, are constantly being developed. Even for commonly used imaging methods, pixel or voxel resolution has increased, as has information density. Clinical CT and MRI, for example, have sub-millimeter spatial resolution, whereas ultrasound has even better spatial resolution while approaching real-time temporal resolution. Medical imaging data is isolated and gathered under unusual conditions. Even though the clinic has a lot of medical imaging data, there isn't a standardized way to acquire it.

2.3 Medical imaging

Measures There is a lot of variety in terms of equipment and scanning settings, which leads to the phenomenon known as "distribution drift." Due to patient confidentiality and clinical information. Architectures of networks. Deep neural networks are more capable of generalization and have a higher model capacity than shallow neural networks. For a single job, deep models trained on large scale annotated datasets yield remarkable results, considerably beyond standard algorithms or even human capacity. Increasing the depth of it. VGGNet (Simonyan & Zisserman, 2015), Inception Net (Szegedy et al., n.d.), and ResNet (He et al., 2016) reflect a research trend that began with AlexNet (KrizhevskyAlex, Ilya Sutskever, 2012) and continued with VGGNet (Simonyan & Zisserman, 2015), Inception Net (Szegedy et al., n.d.), and ResNet (He et al., 2016). As shown in DenseNet (Huang et al., 2017) and U-Net (Weng & Zhu, 2021), using skip connections makes a deep network more trainable. The U-net was suggested initially to deal with segmentation, whereas the other networks were created to deal with picture categorization. Deep supervision (Lee et al., 2015) boosts discriminative capacity even further.

2.4 Transfer Learning

The practice of applying what you've learned from one problem to a different but similar target problem is known as transfer learning (TL). To accelerate training convergence and improve accuracy, one popular TL strategy is to employ a deep network trained on ImageNet and carefully tune it to a medical imaging job. Because there are so many annotated datasets available, such TL techniques have a lot of success. ImageNet, on the other hand, is made out of natural photos, and its pre-trained models are just for 2D images, thus it isn't always the greatest option for medical imaging, especially in small-sample settings (Raghu et al., 2019). Because DL has been utilized in so many different medical imaging applications, it's practically impossible to provide all of the relevant information in a single article. As a result, we'll go over a variety of frequent clinical circumstances, including chest, brain, cardiovascular, abdomen, and microscopic imaging.

2.5 COVID-19 Detection Using AI

Lung disorders are associated with an increased risk of mortality and morbidity. Lung cancer, chronic obstructive pulmonary disease (COPD), pneumonia, and tuberculosis are among the top causes of death throughout the world (TB). COVID-19 presently has the same mortality rate as tuberculosis. Imaging is critical for diagnosis, therapy planning, and understanding more about the causes and processes of this and other lung illnesses. Patients admitted to the hospital are also more likely to develop respiratory issues. As a result, chest radiography is the most often used radiographic method, accounting for more than one-third of all radiology exams. The most often utilized imaging modalities for the chest are plain radiography and computed tomography. CT is excellent for in vivo examination of the lungs due to the considerable contrast in density between air-filled lung parenchyma and tissue, allowing for high-quality and high-resolution pictures to be obtained even at low radiation doses. Oncology patients are diagnosed and staged using nuclear imaging (PET or PET/CT). Although MRI is limited in the lungs, it can provide valuable functional information. Because of the high reflection of sound waves at the air-tissue interface, ultrasound imaging is difficult. Point-of-care ultrasonography, on the other hand, is routinely used to monitor COVID-19 patients, for whom the first deep learning-based decision support apps have already been developed.

COVID-19 screening remains challenging due to the spatial complexity of three-dimensional volumes, the difficulty of identifying infected regions, and the little difference between COVID-19 and other viral pneumonia in chest CT (Multiple, 2020). In (Multiple, 2020), a novel attention-based deep 3D multiple instance learning (AD3D-MIL) method for screening COVID-19 with weak labels is described. The authors obtained 460 chest CT pictures, including 230 for COVID-19, 100 for common pneumonia, and 130 for no pneumonia (healthy people or suffer from other diseases). The total accuracy, AUC, and Cohen kappa score were respectively 97.9%, 99.0%, and 95.7 %. These advantages show that this algorithm can be a useful tool for assisting in the COVID-19 screening process. The total accuracy, AUC, and Cohen kappa score were respectively 97.9%, 99.0%, and 95.7 %. These advantages show that this algorithm can be a useful tool for assisting in the COVID-19 screening process.

SARS-CoV-2 is a virus that causes severe respiratory illnesses. It's an RNA virus that may infect both people and animals (Multiple, 2020). In (Multiple, 2020) Normal, pneumonia, and COVID-19 were the three types of chest pictures used. A DL model was used to detect COVID-19. To restructure the dataset, the fuzzy Color approach was used as a preprocessing step. By mixing structured pictures with original photos, the stacking approach was also used to create a new dataset with improved quality. The stacked dataset was used to train two deep learning models (MobileNetV2 and SqueezeNet). In addition, the Social Mimic optimization approach was used to process the model characteristics. Then, using Support Vector Machines, effective features were combined and classified (SVM). The suggested technique has a 99.27 % overall classification rate, indicating that the model may successfully aid in the detection of COVID-19 disease.

DL methods were employed to distinguish COVID-19 from Influenza-A viral pneumonia in (Butt et al., 2020). Three different CT sample groups were used. The COVID-19 group, which includes 618 CT samples from 110 patients, is the first. 224 CT samples from 224 individuals with Influenza-A viral pneumonia make up the second group. There are 175 CT samples from healthy adults in the third group. Traditional ResNet was used to extract features. COVID-19 was categorized with an overall accuracy of 86.7 % by models using location-attention mechanisms, demonstrating that AI may be a viable tool for diagnosis in the radiology department.

(Imran et al., 2020) has developed an AI-based test for COVID-19 initial diagnosis. A smartphone application named AI4COVID-19 can help distribute the test. To make the diagnosis, the application needs to capture a cough for two seconds. The cough samples are evaluated by a cloud-based AI engine, and the program provides an initial diagnostic in under a minute. Cough, on the other hand, is a common symptom of more than twenty disorders not linked to COVID-19. As a result, diagnosing COVID-19 just based on cough is difficult, and this poses a significant challenge. Three concurrent categorization methods produced by three distinct groups make up the AI engine. An automated mediator cross-validates the classifier's results. If the three classifiers do not agree, the program shows ‘Test inconclusive.' This innovative design improves the diagnostic' validity, making it superior than stand-alone classifiers with a binary diagnostic. The AI engine was shown to be able to distinguish COVID-19 patient cough from a variety of non-COVID-19 coughs with an accuracy of more than 90% (Imran et al., 2020).

A new approach for detecting COVID-19 through integrated smartphone sensors was described in (Maghded et al., 2020). Because all radiologists currently use cellphones for numerous everyday tasks, the concept provides a cost-effective option. Other people can use the framework on their cellphones to detect the infection as well. With powerful CPUs, vast amounts of RAM, and a huge number of sensors, smart phones have attained a high level of performance. The created AI enables the framework to interpret the readings from smart smartphone sensors in order to predict the severity of pneumonia and the disease's fate. The framework is divided into four layers: input sensor measurements layer, sensor setup layer, calculating symptoms disease layer, and foreseeing infection layer using mixed approaches. A transfer learning strategy was also used to improve the ML model in the last stage. The framework is reliable since it relies on multiple readings from many sensors based on the infection's related symptoms.

The use of DL techniques has yielded excellent results in identifying COVID-19 from chest X-rays. COVID-19 was detected from chest X-ray images using an effective convolutional network architecture (Luz et al., 2021). A model with an overall accuracy of 93.9 %, sensitivity of 96.8%, and positive prediction of 100 % has been developed. (Abbasian et al., 2020) employed AI to diagnose COVID-19 in a quick and effective manner. A total of two sets of CT slices were used. The first group is the COVID-19 group, which includes CT slices from 108 individuals whose infection was confirmed in the lab. CT slices for 86 individuals with different atypical and viral pneumonia illnesses are included in the non-COVID-19 group. This study made use of a total of 1020 CT slices. AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, and Xception are ten convolutional neural networks (CNNs) that were used to provide an overview of the role of AI in COVID-19 disease diagnosis. The results demonstrated that DL is capable of detecting COVID-19 and other illnesses with high accuracy. ResNet-101 and Xception produced the best results. ResNet-101 distinguished between COVID-19 and non-COVID-19 cases with an AUC of 0.994. (sensitivity, 100 % specificity, 99.02 % , accuracy, 99.51 %). Xception had an AUC of 0.994. (sensitivity, 98.04 % specificity, 100 % , accuracy, 99.02 %). Radiologists, on the other hand, did well, with an AUC of 0.873. (sensitivity, 89.21 % , specificity, 83.33 % , accuracy, 86.27 %). For detecting COVID-19 infection, the ResNet-101 model provides the highest sensitivity.

COVID-19 was detected from X-ray images using a CNN-based transfer learning algorithm in (Chowdhury et al., 2020). For COVID-19, viral pneumonia, and normal chest X-ray pictures, the dataset used was made up of 190, 1345, and 1341 pictures, respectively. A training set of around 2600 photos of each class was used to train four distinct CNNs (AlexNet, ResNet18, DenseNet201, and SqueezeNet). The training dataset was created using the image augmentation approach. The classification of two separate methods was investigated using CNN networks (normal and COVID-19 pneumonia; normal, viral, and COVID-19 pneumonia). SqueezeNet outperforms the other three deepCNN networks, according to the results. For both systems, the classification accuracy, sensitivity, specificity, and precision were 98.3%, 96.7%, 100%, 100% and 98.3%, 96.7%, 99%, 10%, respectively. SqueezeNet was shown to be able to detect COVID-19 from X-ray pictures and discriminate it from normal and viral pneumonia pictures in this study.

To categorize X-ray pictures into normal, pneumonia, and COVID-19, a deep CNN based on the concatenation of Xception and ResNet50V2 networks was suggested in (Rahimzadeh & Attar, 2020a). The collection includes 180 COVID-19 photos, 6054 pneumonia chest pictures, and 8851 x-ray pictures of healthy persons. A new strategy for evaluating the neural network (NN) when the dataset is not balanced was developed due to the low amount of COVID-19 X-ray pictures. The training set was divided into eight phases, with 633 pictures utilized in each level. The COVID-19 class earned an average accuracy of 99.50 %, with an overall accuracy of 91.4 %. As a result, this well-trained network will be beneficial in medical diagnostics. COVID-19 was detected from chest X-ray images using a deep CNN-based model (COVID-Net) described in (Multiple, 2020). There are 13975 chest X-ray pictures in the collection, spread among 13870 patient cases. The planned COVID-Net has been tested and found to be 93.3 % accurate. The early detection and categorization of COVID-19 illness is critical for disease management and therapy. CT imaging is the quickest and most accurate way to diagnose COVID-19 illness. Because CT equipment are readily available in hospitals, they can be used to classify COVID-19 patients early. However, the specialist's time is crucial, especially when the disease is fast spreading. As a result, an automated approach for analyzing CT scans must be developed in order to save specialists' time (Singh et al., 2020). A multi-objective differential evolution (MODE)-based CNN was used in (Singh et al., 2020) to classify people as positive (infected) or negative (uninfected) (not infected). In terms of accuracy, F-measure, sensitivity, specificity, and Kappa statistics, the proposed model outperforms competitive models such as artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS), and CNN models by 1.9789%, 2.0928%, 1.8262%, 1.6827%, and 1.927%, respectively. As a result, the suggested approach can accurately identify chest CT images.

Truncated Inception Net, a DL-based CNN model, was suggested in (Iwendi et al., 2020). The algorithm has a 99.96 % accuracy in classifying COVID-19 positive patients among pneumonia and healthy subjects. It also classified X-ray images of COVID-19 positive cases among pneumonia and TB cases with 99.92 % accuracy. A new DL-based approach for detecting COVID-19 using raw chest X-ray images was proposed in (Aboughazala & Mohammed, 2020). With a classification accuracy of 98.08 % for binary classification (COVID vs. No-Findings) and 87.02 % for multi-class classification (COVID vs. No-Findings vs. Pneumonia), the model has attained a correct diagnosis. COVID-19 was detected on chest X-ray images using DL-based CNN models in (Ghaderzadeh & Asadi, 2021). Data augmentation was employed in this project. A total of 6432 X-ray pictures for Covid-19 afflicted, normal, and pneumonia patients were obtained to assess performance. The dataset was split into 5467 training images and 965 validation images. Three models, Inception V3, Xception, and ResNeXt, were compared in terms of accuracy. The results showed that the Xception model surpasses the others and achieves the best level of accuracy (97.97 %). (M. Z. Islam et al., 2020) suggested a DL method for diagnosing COVID-19 from X-ray images by combining CNN and LSTM. The purpose of CNN is to extract deep characteristics. The detection method, on the other hand, was carried out using

LSTM. The collection consists of 4575 X-ray pictures, 1525 of which are COVID-19 pictures. The accuracy reached was 99.4%.

2.6 Employing ML in prediction

Researchers have used machine learning approaches to forecast a variety of ailments, including coronary artery disease and cardiovascular disease. These prediction algorithms are extremely helpful in effectively managing these diseases (Rustam et al., 2020). (Rustam et al., 2020) shows how machine learning algorithms can predict the amount of future COVID-19-affected patients. Linear regression (LR), least absolute shrinkage and selection operator (LASSO), SVM, and exponential smoothing (ES) were used as standard predictive models. Each model makes three major predictions in the following ten days: the number of new confirmed cases, the number of fatalities, and the number of recoveries. The data show that using these tactics to combat the present COVID-19 epidemic is a viable option. It was found that the ES performance was the best among all the utilized models followed by LR and LASSO and it has performed well in predicting the novel affirmed cases, death rate as well as recovery rate, while SVM performance was the worst in all the prediction scenarios given the accessible dataset.

2.7 Segmentation from CT Images

CT scans are preferred over X-rays because they provide a three-dimensional image of the lungs. Some findings on CT scans, such as GGO in the early stages and pulmonary consolidation in the later stages, suggest infection (Fan et al., 2020)(Ai et al., 2020). Automatic diagnosis of infected regions in the lung using CT scans has a lot of promise for assisting radiologists in COVID-19 detection. However, due to the large disparity in infection characteristics and low-intensity contrast between infections and normal tissues, segmenting these impacted regions remains a difficult endeavor. Furthermore, collecting a huge amount of data in a short period of time is impracticable. As a result, training a deep model will be impossible. To address these issues, the authors suggested a novel deep Network to segregate COVID-19 Lung Infected from chest CT slices in (Fan et al., 2020). A parallel partial decoder was used to collect a high-level of attributes and create a global map. For modeling the boundaries and enhancing the representations, implicit reverse attention and explicit edge attention were also used. A semisupervised segmentation system based on a randomly selected propagation mechanism has also been described to address the absence of labeled data. Only a few identified photos and the use of largely unlabeled data are required in this system. Additionally, learning capability is improved, and greater performance may be achieved. This study concentrates on the task of segmenting diseased lungs, but clinically, it is necessary to first categorize COVID-19 patients. Then, for further treatment, segmentation for contaminated areas might be done.

With the growth in the number of suspected cases, manual segmentation of CT scans has become a tough process, necessitating the development of a reliable and automated segmentation system. Current image segmentation approaches are unable to achieve satisfactory performance due to the change of imaging characteristics of COVID-19 and their resemblance to the background (Yan et al., 2020). COVID-19 infection regions and the whole lung were segmented from chest CT images using a new deep CNN (COVIDSegNet) in (Yan et al., 2020).A new chest CT image collection was created, consisting of 21,658 annotated chest CT scans from 861 individuals who had been diagnosed with COVID-19. The Feature Variation (FV) block has been offered as a solution to the problem of COVID-19 pneumonia being difficult to distinguish from the lung. In addition, Progressive Atrous Spatial Pyramid Pooling (PASPP) has been used, which gradually increases the size of the pyramid.

2.8 Challenges of Using ML for COVID-19 Detection

A huge amount of data is necessary to construct effective machine learning systems. Using machine learning for COVID-19 research is currently fraught with difficulties. The absence of standard data is one of the key obstacles of utilizing DL to diagnose COVID-19 (M. M. Islam et al., 2021)(Alafif et al., 2021). Another difficult problem is the sample imbalance in the dataset. In comparison to pneumonia and normal cases samples, COVID-19 samples had a modest number of X-ray and CT scans. Data augmentation is the most often used method for dealing with an unbalanced dataset. This method uses flipping, rotation, zooming, random noise addition, and other techniques to create new lesions from COVID-19 data. Another advantage of this method is that it can eliminate overfitting problems (Toğaçar et al., 2020)(Butt et al., 2020)(Luz et al., 2021)(Chowdhury et al., 2020)(Rahimzadeh & Attar, 2020a)(L. Wang et al., 2020)(Ozturk et al., 2020) (M. M. Islam et al., 2021) (Jain et al., 2021)(Pratap et al., 2020).

CHAPTER 3

DEEP LEARNING BASED FEATURE EXTRACTION COUPLED WITH MULTI CLASS SVM FOR COVID-19 DETECTION

3.1 Introduction

In this study, two classifiers SVM and KNN were employed to classify COVID-19, common pneumonia and healthy individuals CT scan images, before training the classifiers, handcrafted LBP features and Automatic deep learning features of seven pre-trained networks were extracted, the training of the classifiers was conducted on the extracted features, to improve the performance of the classifiers, a new feature was proposed by concatenating the LBP and the CNN features to train the classifiers, this proposed feature shows improvement in the performance of the classifiers compared to the performance of the classifiers on LBP or CNN features. The results achieved by the proposed model are accuracy of 99.4%, a sensitivity of 99.3%, specificity of 99.3% and AUC of 99.3%. This study will help in detecting the COVID-19 at an early stage in order to avoid the spread of the virus and crippling the economy

3.2 COVID-19 DETECTION

This section describes the characteristics of the dataset used, the proposed feature extractions techniques and the machine learning models for COVID-19 detection

3.2.1 Datasets

In this study, three datasets were used in classifying CT scan images. The collected datasets include three classes namely: COVID-19 positive, healthy individuals and common pneumonia. The first dataset (Yang et al., n.d.) contains CT scan images of 349 COVID-19 positives and 397 healthy individuals, the second dataset (Tao Yan, Pak Kin Wong, Hao Ren c, Huaqiao Wang, Jiangtao Wang, n.d.) contains CT scan images of 328 common pneumonia and 371 COVID-19 positive patients and the third dataset (Soares et al., 2020) contains scan images of 1252 COVID-19 positive class and 1229 healthy individuals. The overall dataset comprises

1608 healthy individuals, 1972 COVID-19 positive and 328 patients with common pneumonia as presented in Table 3.1.

Ta**ble 3.1:** Compiled dataset

Dataset	COVID-19 Positive	Common Pneumonia	Healthy Individuals
(Yang et al., n.d.)	349	NA	397
(Tao Yan, Pak Kin Wong, Hao Ren c, Huaqiao Wang, Jiangtao Wang, n.d.)	371	328	NA
(Soares et al., 2020)	1252	NA	1229
Total Number of CT scan images per class	1972	328	1608

Figure 3.1: COVID-19 positive CT image

Figure 3.2: COVID-19 Healthy individuals CT image

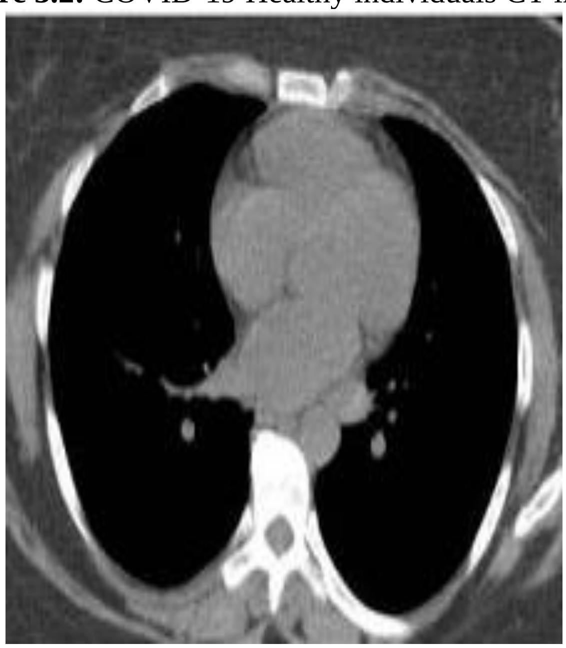


Figure 3.3: Common Pneumonia CT Image

3.2.2 SVM

The SVM (W. C. Wang et al., 2013) When a training set is provided, the algorithm generates a hyperplane that optimizes the margin between two input classes, which is referred to as classification. For example, if two classes of data are linearly separated, the system can have several hyperplanes separating them. SMV finds the most ideal hyperplane with the largest margin among all accessible hyperplanes, where the margin is the difference in distance between the hyperplane and the support vectors. Given a set of training data $\{(x_i, d_i)\}_i^N$ (d_i is the actual value, x_i represents the input vector and N is the data number), given that the SVM function is:

$$y = f(x) = w\phi(x_i) + b \qquad (1)$$

where $\phi(X)$ is mapped non-linearly from input vector x, which are input feature spaces.
Then, the SVM equation is given as (W. C. Wang et al., 2013) :

$$f(x, \alpha_{j,} \alpha_{i^*}) = \sum_i^N (\alpha_i - \alpha_{i^*}) K(x, x_i) + b \qquad (2)$$

$k(x_i, x_j)$ After conducting nonlinear mapping, is the kernel function in the feature space, and b is the bias term. The Gaussian Radial Basis Function (RBF) is the most widely utilized kernel function because it outperforms linear and polynomial kernels by not only being able to transfer non-linearly training data into infinite-dimensional space but also being easy to construct (W. C. Wang et al., 2013) and it is given as:

$$k(x_1, x_2) = exp(-\gamma \|x_1 - x_2\|^2) \qquad (3)$$

where γ is the kernel parameter

3.2.3 K-NN

K – Nearest Neighbors (k-NN) is a non-parametric method of classification (Altman, 1992) when classifying using k-NN, the entity to be ranked is decided upon by its neighbours and assigned to the most comparable class of its closest neighbours. Three-class k-NNs is used as part of the research.

3.2.4 Local Binary Pattern

The centre pixel (xc,yc), the ordered binary set designated as LBP is derived by comparing the grey value of the centre pixel (xc,yc) with the pixels of its eight neighbours. As a result, the LBP code is a decimalized version of an octet binary integer:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} S(i_n - i_c)2^n$$

Where ic denotes the grey value of the center pixel (xc,yc), as well as the grey value of its eight neighbors' pixels. Any monotonous modification of the grey level has no effect on the LBP code, and the local binary code stays unaltered after transformation.

$$S(i_n - i_c) = \begin{cases} 1 i_n - i_c \geq 0 \\ 0 i_n - i_c < 0 \end{cases}$$

3.2.5 GoogleNet

GoogleNet is a 22 layer network comprising of the input layer, convolution layers, max-pooling and softmax classifier, the main things that make the GoogleNet different is the 1 × 1 convolution, network in network and the global average pooling. GoogleNet won the ILSVRC 2014 competition with a low error rate compared to VGG(Szegedy et al., n.d.).

3.2.6 Convolutional Neural Network

A CNN is a deep model in which convolutional filters and pooling operations are done alternatively on the immediate neighbourhoods of each pixel in the raw input, resulting in complicated high-level features (Li et al., 2017). CNNs have largely been used on 2D pictures and have shown to improve image classification ability. Convolution and pooling techniques that are effective offer discriminative CNN features that are used by different filters. CNN creates a huge number of feature maps when it employs pretrained models. It's always been challenging to choose the right collection of features for machine learning algorithms (Simon et al., 2020).

3.2.7 Transfer Learning

Transfer learning is a research topic of machine learning. It focuses on preserving information learned when solving a particular problem and adapting it to a particular but connected problem. (Apostolopoulos & Mpesiana, 2020; Haque & Rahman, 2020; Hussein et al., 2019; Learning, 2020; Mahmud et al., 2020). In training the pre-trained network for another problem, some features of the pre-trained models may be modified, such modifications are layers to freeze, layers to be inserted, and some hyperparameter values changed

3.2.8 ResNet

ResNet or Residual Network (He & Sun, 2016) It's a deep learning algorithm used to classify images. The core principle behind ResNet is to deal with vanishing gradients that degrade network performance caused by accumulating a convolution layer over a pooling layer in deep network architecture, shortcuts that provide identification is a residual block, the notion of adding skip connections essentially eliminates a high training error, other deep networks do not include an identity link, which is why ResNet may not. The input layer accepts a 224 × 224 image size.

3.2.9 GoogleNet

GoogleNet is a 22-layer network composed of the input layer, convolution layers, max-pooling and softmax classifier, the key feature that makes GoogleNet different is the 1×1 convolution, networking and global average pooling. GoogleNet won the 2014 ILSVRC competition at a low error rate relative to VGG. (Szegedy et al., n.d.).

3.2.10 ShuffleNet

ShuffleNet (X. Zhang et al., n.d.) is a 50-layer deep learning model that uses AlexNet's group convolution on the first convolution layer. Although group convolution greatly decreases computation, one downside is that the performance of such channels is powered by a small fraction of the input. To resolve this problem, to address this issue, the channels which are also differentiable are shuffled in ShuffleNet to address this issue.

3.2.11 VGG

VGG (Zisserman Karen, 2015) VGG16 has 16 layers, 5 convolutional layers, 3 trainable layers, and the remaining layers are max-pooling layers, while VGG19 has 19 layers. This design came in second place in the 2014 Visual Recognition Competition. i.e. ILSVRC-2014

.

3.3 Training

In this study three datasets were merged to detect COVID-19, the three datasets contained three classes, COVID-19, common pneumonia and healthy individuals. Before training, the data was preprocessed by resizing the images to 224 by 224, also, several data augmentation such as random reflection, random rotation, random rescale, random translation along X-axis and random translation along Y-axis was performed to increase the number of images, improves training and to reduce overfitting (Ghassemi et al., 2020; Loey et al., 2020). After the augmentation, two features extraction techniques handcrafted Local Binary Pattern and automatic deep features from seven pre-trained models MobileNetv2, GoogleNet, ResNet-50, ResNet-101, ShuffleNet, VGG16 and VGG19 were extracted. 80% of the data was used for training and 20% for testing.

The training for this study is carried out in stages using two classifiers, KNN and SVM, in the first stage, textual features were extracted using LBP and were classified by KNN and SVM, in the second stage, high-level deep features were extracted from each of the seven pre-trained models MobileNetv2, GoogleNet, ResNet-50, ResNet-101, ShuffleNet, VGG16 and VGG19, the features of the last pooling layer of each network were used for the training of the two classifiers KNN and SVM, in the third stage the LBP features and features extracted from each of the seven pre-trained networks are concatenated then classified using the two classifiers KNN and SVM, the concatenation of the two features will give the advantage of textual and high-level deep features of the CT scan images. Figure 3.4. Shows the classification process for the COVID-19 detection.

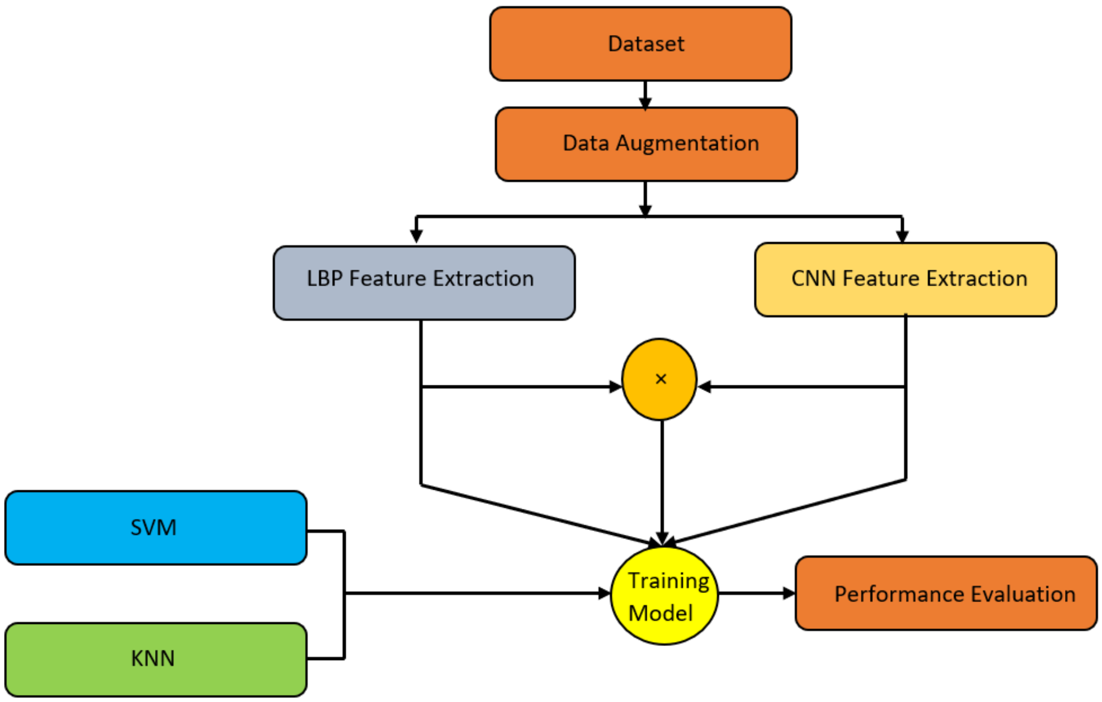


Figure.3.4: Shows the classification process for the COVID-19 detection.

3.4 Results and Discussion

Within the scope of the study, three classes of CT scan images COVID-19, common pneumonia and healthy individuals classes were classified using two classifiers, KNN and SVM. Before training the classifiers, handcrafted LBP features and deep high-level features of seven deep learning models were extracted. In Table 3.2, SVM and KNN classifiers performance were compared to determine the best performing model with LBP features as inputs, the SVM achieves an accuracy of 97.5% sensitivity of 98.7% a specificity of 96.1%, F1 score of 97.5%, the precision of 96.4%, Yonden index of 94.8% and AUC of 97.4%. The SVM as shown in Figure 3.4 outperformed the KNN in terms of accuracy, sensitivity, specificity, F1 score, precision, Yonden index and AUC. This shows that the SVM can efficiently detect COVID-19, common pneumonia and healthy individuals CT scan images based on LBP features extracted from the CT scan images.

Table 3.2: SVM and KNN performance on LBP extracted features

Models	Accuracy (%)	Sensitivity(%)	Specificity(%)	F1 Score(%)	Precision (%)	Yonden Index(%)	AUC(%)
LBP SVM	97.5	98.7	96.1	97.5	96.4	94.8	97.4
LBP KNN	91.8	92.5	90.9	92.0	91.5	83.4	91.7

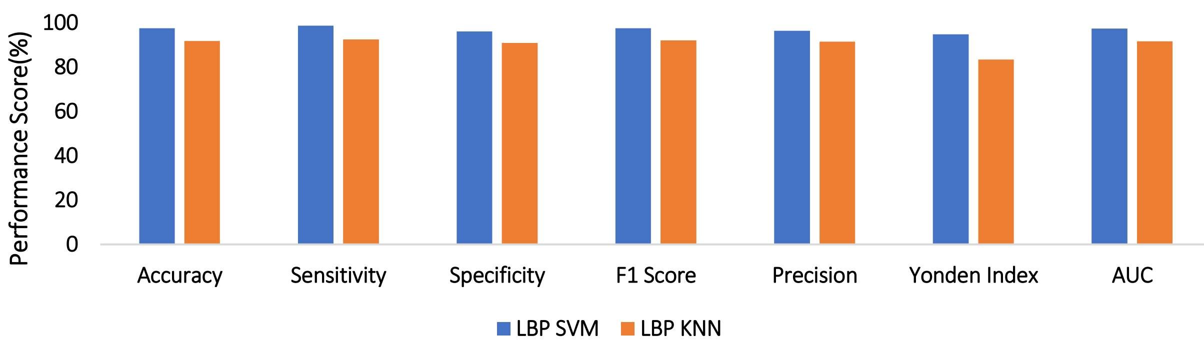


Figure.3.5: Performance comparison between SVM and KNN on LBP extracted features.

In Table 3.3 and Table 3.4, SVM and KNN were used to classify seven deep learning models extracted features respectively, in Table 3.3, SVM with extracted features of VGG-19 achieves an accuracy of 98.7, the sensitivity of 98.8%, specificity of 98.7%, F1 score of 98.8%, the precision of 97.5%, Yonden index of 97.5% and AUC of 98.7%. In table 3.4, the SVM turns out to achieve the highest performance with features extracted from VGG-19. The KNN classifier achieved the highest performance with features extracted from VGG-16, as presented in table 4, the KNN with VGG-16 extracted features achieves an accuracy of 96.2%, the sensitivity of 96.9%, a specificity of 95.5%, F1 score of 96.3%, precision of 95.7%, Yonden

Index of 92.4% and AUC of 96.2%. If we compare the performances of the SVM and KNN on both the LBP and the deep learning features, it will be observed in Figure 3.6 and Figure 3.7 that the SVM shows superiority in terms of classifying the CT scan images for the three classes.

Table 3.3: CNN extracted features with SVM classifier models performance

Models	Accuracy(%)	Sensitivity(%)	Specificity(%)	F1 Score(%)	Precision(%)	Yonden Index(%)	AUC(%)
ResNet-50	95.9	95.6	95.1	95.9	96.3	90.7	95.4
ResNet-101	95.9	95.6	96.1	95.9	96.3	91.7	95.9
VGG-16	97.5	97.4	97.5	97.5	97.5	94.9	97.5
VGG-19	98.7	98.8	98.7	98.8	98.8	97.5	98.7
GoogLeNet	97.8	98.1	97.4	98.4	98.8	95.5	97.8
MobileNetv2	97.8	97.5	98	97.8	98.14	95.5	97.7
ShuffleNet	93.4	89	97.4	93.0	97.3	86.4	93.2

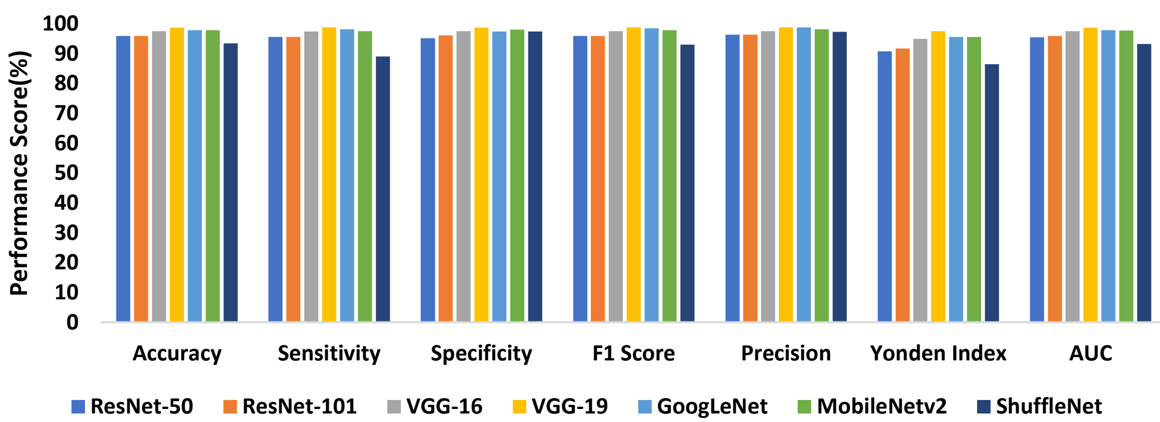


Figure.3.6: CNN extracted features with SVM classifier models performance

Table 3.4: CNN extracted features with KNN classifier models performance

Models	Accuracy(%)	Sensitivity(%)	Specificity(%)	F1 Score(%)	Precision(%)	Yonden Index(%)	AUC(%)
ResNet-50	94.3	92.5	96.1	94.3	96.2	88.6	94.3

ResNet-101	94.9	96.3	93.5	95.1	94.0	89.8	94.9
VGG-16	96.2	96.9	95.5	96.3	95.7	92.4	96.2
VGG-19	93.1	95	90	93.0	91.2	85	92.5
GoogLeNet	90.8	91.3	89.3	90.8	90.2	80.6	90.3
MobileNetv2	93.1	95.6	90.3	93.3	91.2	85.9	93.0
ShuffleNet	90.9	91.9	98.7	91.1	90.3	90.6	95.3

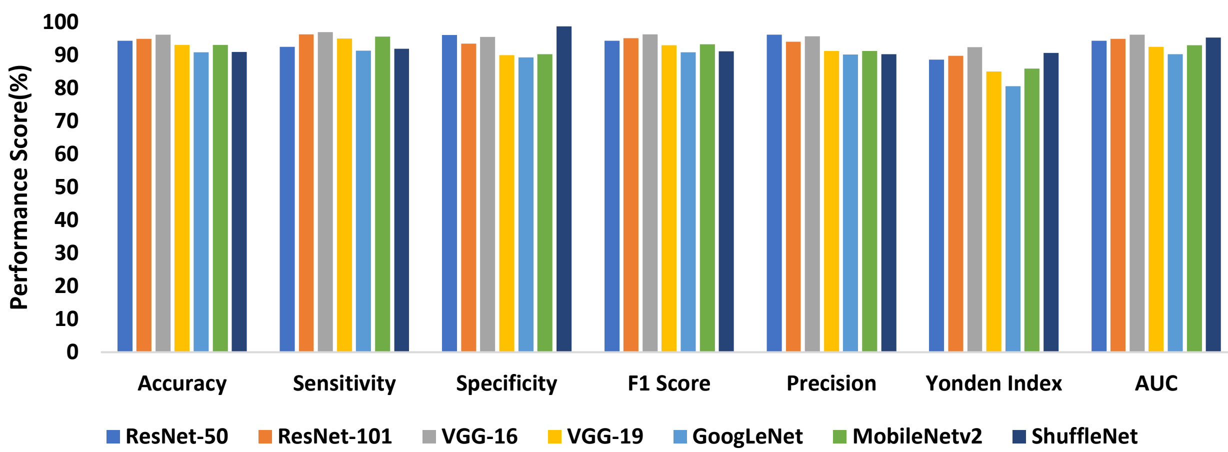


Figure 3.7: CNN extracted features with KNN classifier models performance

Table 3.5: CNN and LBP extracted features with SVM classifier models performance

Models	Accuracy(%)	Sensitivity(%)	Specificity(%)	F1 Score(%)	Precision(%)	Yonden Index(%)	AUC(%)
ResNet-50+ LBP	98.7	97.5	100	98.1	98.8	97.5	98.8
ResNet-101+ LBP	94.3	92.5	96.1	94.3	96.2	88.6	94.3
VGG-16+ LBP	97.5	100	94.8	97.6	95.3	94.8	97.4
VGG-19+ LBP	98.1	98.7	97.4	98.1	97.6	96.1	98.05

GoogLeNet+ LBP	98.1	97.5	98.7	98.1	98.8	96.2	98.1
MobileNetv2+ LBP	98.7	100	97.4	98.8	97.6	97.4	98.7
ShuffleNet+ LBP	96.2	93.8	97	93.2	92.7	90.8	95.4

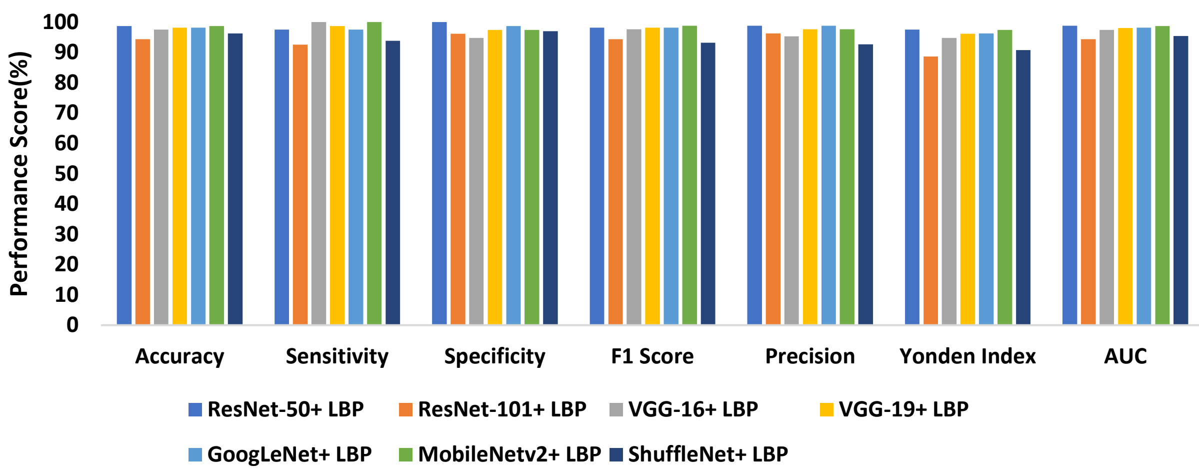


Figure 3.8: CNN and LBP extracted features with SVM classifier models performance.

Table 3.6: CNN and LBP extracted features with KNN classifier models performance

Models	Accuracy(%)	Sensitivity(%)	Specificity(%)	F1 Score(%)	Precision(%)	Yonden Index(%)	AUC(%)
ResNet-50+ LBP	99.1	99.3	98.7	98.7	98.2	98	99
ResNet-101+ LBP	97.2	98.7	95.5	97.2	95.8	94.2	97.1
VGG-16 + LBP	98.1	97.5	98.7	98.1	98.8	96.2	98.1
VGG-19 + LBP	99.4	99.3	99.3	99.0	98.8	98.6	99.3

GoogLeNet + LBP	98.4	98.7	98	98.4	98.2	96.7	98.4
MobileNetv2+ LBP	97.2	98.7	95.5	97.2	95.8	94.2	97.1
ShuffleNet+ LBP	95.9	98.5	93.1	96.2	94.1	91.6	95.8

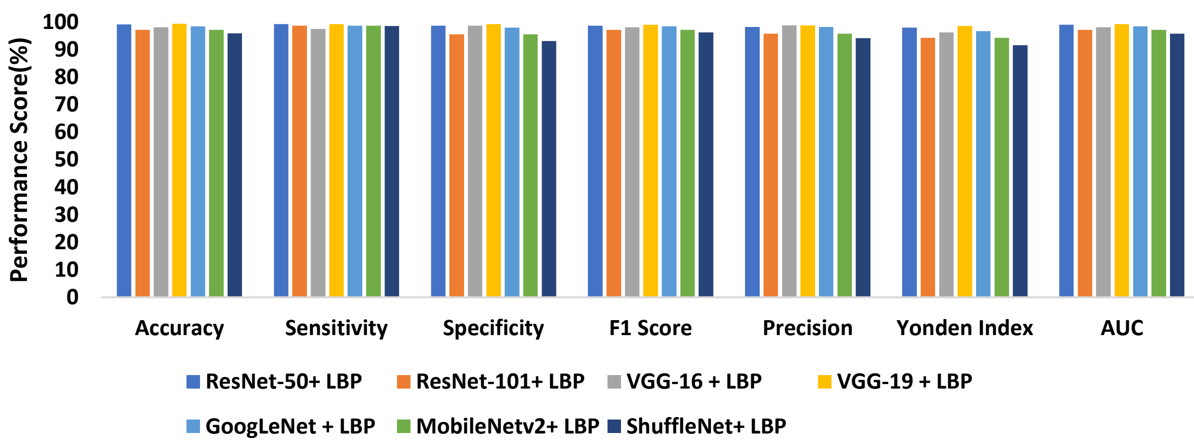


Figure 3.9: CNN and LBP extracted features with KNN classifier models performance

To take the advantage of both the handcrafted LBP features and the automatic deep learning features, in this study, the LBP and CNN features were extracted and then concatenated, the KNN and the SVM classifiers were trained using these concatenated features. In Table3.5 and Table 3.6, the results of the SVM and KNN classifiers were presented respectively. In table 10, the SVM achieved the highest performance in terms of accuracy with ResNet-50+ LBP and MobileNetv2+ LBP features with an accuracy of 98.7%, in terms of sensitivity VGG-16+ LBP and MobileNetv2+ LBP achieved a sensitivity of 100% each, for the specificity ResNet-50+ LBP achieves the highest specificity of 100%. In table 3.6, the KNN classifier trained with VGG-19 + LBP features shows superiority in terms of performance as it achieves an accuracy of 99.4%, a sensitivity of 99.3%, a specificity of 99.3%, F1 score of 99%, precision of 98.8%, Yonden Index of 98.6% and AUC of 99.3. This is because the VGG-19 pretrained model can deal with the effect of overfitting. Based on Table 3.2 and Table 3.3, the concatenated features have improved the detection of COVID-19 on CT scan images, as clearly shown in Figure 3.8 and Figure 3.9, the two classifiers can effectively classify the new proposed concatenated features. The variation in the performance of the classifiers is due to the different architecture of the deep learning models, each deep learning has its advantages when extracting features. In Table 3.7, compared to the state of the art models in which three classes of COVID-19, common pneumonia and Healthy individuals CT scan images were detected, our models show greater performance in terms of accuracy, sensitivity specificity and AUC.

Table 3.7: State of the art Model comparison with the proposed model

Ref.	Models	Accuracy(%)	Sensitivity	Specificity	AUC
(Amyar et al., 2020)	T1 & T2 & T3 512 × 512	91.13	0.94	0.85	0.94
	T1 & T2 & T3 256 x 256	94.67	0.96	0.92	0.97
Proposed Model	VGG-19 + LBP	99.4	99.3	99.3	99.3

It is important to improve COVID-19 detection, particularly on CT scan images that provide slice-level information of the chest. X-ray images were used in many experiments on medical imaging, but CT images were used in only a handful of studies. Hence, we were motivated to carry out this study based on several factors, including insufficient CT scan images data, and the ability to distinguish between common pneumonia, COVID-19 and healthy individuals' CT scan images

CHAPTER 4
EFFICIENT COVID-19 DETECTION USING LOCAL BINARY PATTERN AND DEEP LEARNING FEATURE FUSION IN IoT ERA

4.1.Introduction

Seven pre-trained models AlexNet, GoogleNet, ResNet-50, ResNet-101, ShuffleNet, VGG16 and VGG19 were employed in this study as feature extractors and a multi-class SVM as the classifier to classify COVID-19, Common Pneumonia and Healthy Individual CT scan images, detecting COVID-19 at an early stage are very important, it will reduce the spread of the virus and putting the patient on the right diagnosis, that is why in this study, we improved the performance of deep learning models by augmenting the images and changing the classifier of the pre-trained models with a classifier that is efficient and takes less time to train. With this approach, we are confident that COVID-19 can be detected at an early stage with high accuracy and reliability.

4.2.COVID-19 DETECTION

This section describes the characteristics of the dataset used, the proposed feature extractions techniques and the machine learning models for COVID-19 detection

4.2.1. Datasets

In this research, three datasets have been used to classify CT scan images. The datasets obtained provide three classes: positive COVID-19, healthy individuals and common pneumonia. The first collection of data (Yang et al., n.d.), contains 349 COVID-19 positive and 397 healthy individuals CT scan images, collection of data (Tao Yan, Pak Kin Wong, Hao Ren c, Huaqiao Wang, Jiangtao Wang, n.d.) contains 328 common pneumonia and 371 COVID-19 positive CT scan images and the third collection of data (Soares et al., 2020) contains 1252 COVID-19 Positive class and 1229 Healthy individuals. Meanwhile, to generalized the dataset the three datasets were merged. The overall dataset comprises 1608 healthy individuals, 1972 COVID-19 positive and 328 patients with common pneumonia.

Table 4.1. Compiled dataset

Dataset	COVID-19 Positive	Common Pneumonia	Healthy Individuals
(Yang et al., n.d.)	349	NA	397
(Tao Yan, Pak Kin Wong, Hao Ren c, Huaqiao Wang, Jiangtao Wang, n.d.)	371	328	NA
(Soares et al., 2020)	1252	NA	1229
Total Number of CT scan images per class	1972	328	1608

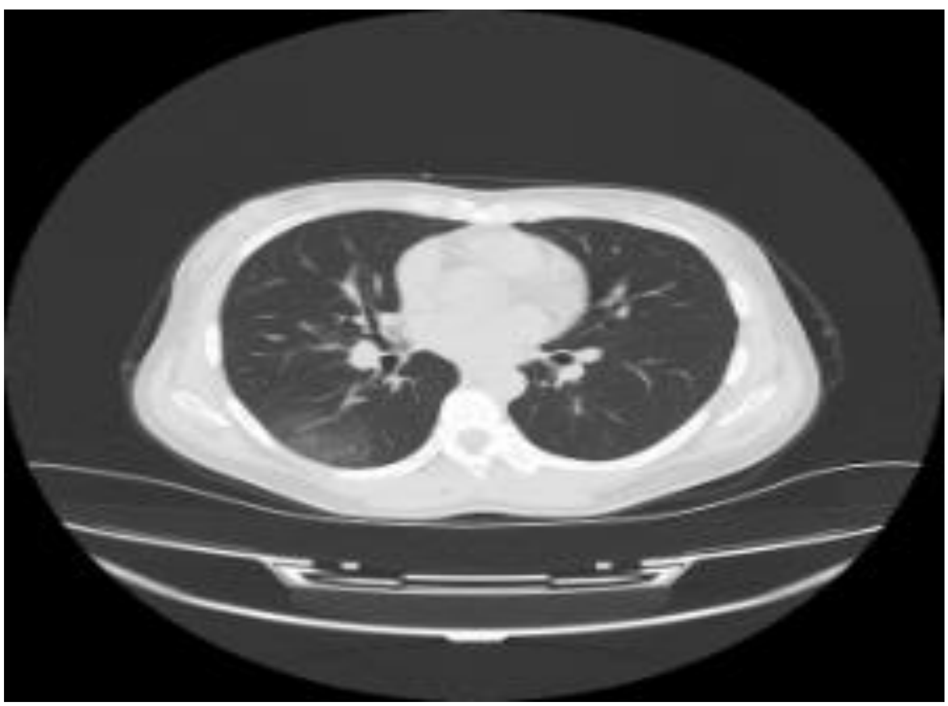


Figure 4.1: COVID-19 positive CT image

Figure 4.2: COVID-19 Healthy individuals CT Image

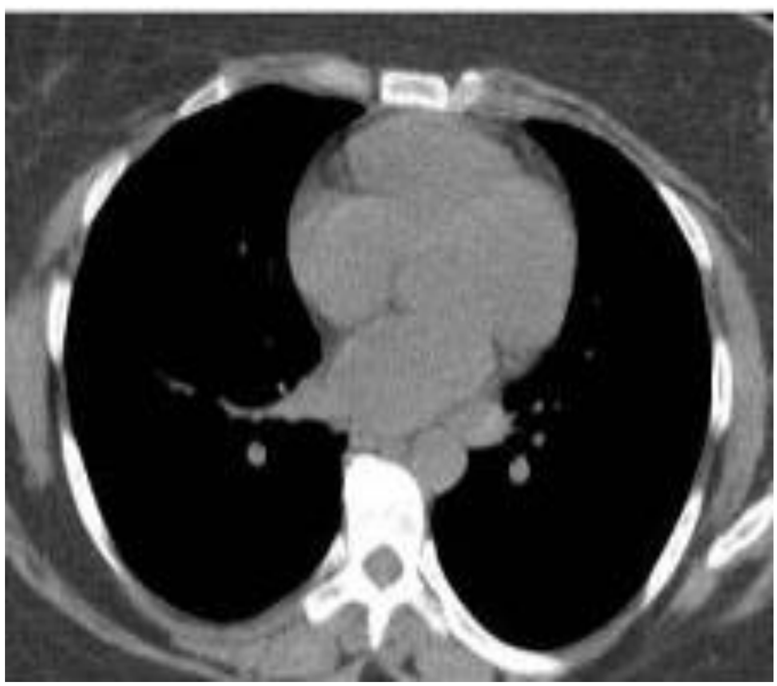


Figure 4.3: Common Pneumonia CT Image

4.2.2. Transfer Learning

This Transfer learning is a research problem in machine learning. It focuses on storing knowledge gained while solving one problem and applying it to a different but related problem (Apostolopoulos & Mpesiana, 2020; Haque & Rahman, 2020; Hussein et al., 2019; Learning, 2020; Mahmud et al., 2020). In training the pre-trained network for another problem, some features of the pre-trained models will be changed, such changes are layers to be frozen, layers to be added and change of some hyperparameters values

4.2.3. ResNet

ResNet or Residual Network (He & Sun, 2016) is a deep learning algorithm used in classifying images. The key concept behind ResNet is to deal with disappearing gradients that degrade network output induced by piling up a convolution layer over a pooling layer in deep network architecture, shortcuts that include identity is a residual block, the idea of inserting skip connections effectively removes a high training error, other deep networks do not contain an identity connection that is why ResNet is different, the ResNet-50 has 50 layers while ResNet-101 has 101 layers. The input layer accepts an image of size 224 × 224.

4.2.4. AlexNet

AlexNet, (KrizhevskyAlex, Ilya Sutskever, 2012) is a deep learning model which consist of five convolution layers, three fully connected layers and 3 max-pooling layers, the AlexNet was first to win the ImageNet challenge in 2012, Rectified Linear Unit was first introduced in AlexNet, this makes it to train faster compared to CNN with tahn function. The AlexNet input layer admits images with the size 227 × 227

4.2.5. GoogleNet

GoogleNet is a 22 layer network comprising of the input layer, convolution layers, max-pooling and softmax classifier, the main things that make the GoogleNet different is the 1 × 1 convolution, network in network and the global average pooling. GoogleNet won the ILSVRC 2014 competition with a low error rate compared to VGG(Szegedy et al., n.d.).

4.2.6. ShuffleNet

ShuffleNet (X. Zhang et al., n.d.) is a deep learning model that has 50 layers, it utilizes the group convolution from AlexNet on 1× 1 convolution layer, the group convolution reduces computation significantly, but the drawbacks of the group convolutions is that the output of certain channels is driven from a small fraction of the input, to address this issue, the channels which are also differentiable are shuffled in ShuffleNet to address this issue.

4.2.7. VGG

VGG (Zisserman Karen, 2015) is a deep learning model which utilizes small filters and max pooling after every convolution, the VGG16 which contains 16 layers, out of the 16 layers 5 are convolutional layers, 3 trainable layers and the remaining layers are max-pooling layers, while the VGG19 has 19 layers. This architecture was the 1st runner up of the Visual Recognition Challenge of 2014 i.e. ILSVRC-2014

4.2.8. SVM

The SVM (W. C. Wang et al., 2013) When a training set is provided, the algorithm generates a hyperplane that optimizes the margin between two input classes, which is referred to as classification. For example, if two classes of data are linearly separated, the system can have several hyperplanes separating them. SMV finds the most ideal hyperplane with the largest margin among all accessible hyperplanes, where the margin is the difference in distance between the hyperplane and the support vectors. Given a set of training data $\{(x_i, d_i)\}_i^N$ (d_i is the actual value, x_i represents the input vector and N is the data number), given that the SVM function is:

$$y = f(x) = w\phi(x_i) + b \qquad (1)$$

where $\phi(X)$ is mapped non-linearly from input vector x, which are input feature spaces.
Then, the SVM equation is given as (W. C. Wang et al., 2013) :

$$f(x, \alpha_{j,} \alpha_{i^*}) = \sum_{i}^{N} (\alpha_i - \alpha_{i^*}) K(x, x_i) + b \qquad (2)$$

$k(x_i, x_j)$ After conducting nonlinear mapping, is the kernel function in the feature space, and b is the bias term. The Gaussian Radial Basis Function (RBF) is the most widely utilized kernel function because it outperforms linear and polynomial kernels by not only being able to transfer non-linearly training data into infinite-dimensional space but also being easy to construct (W. C. Wang et al., 2013) and it is given as:

$$k(x_1, x_2) = exp(-\gamma \|x_1 - x_2\|^2) \qquad (3)$$

where γ is the kernel parameter.

4.3.Data Pre-processing and Training

Data preprocessing in deep learning is the process of perfecting data in such a way that it can fit the input of a network and also, clean and increase the number of the dataset for robust and better training. There are several types of data preprocessing such as resizing, augmentation and smoothing in training medical images.

4.3.1. Data Augmentation

Data augmentation is a method of obtaining additional data from the initial data set, an increase in data increases training performance and prevents over-fitting (Ghassemi et al., 2020; Loey et al., 2020). Several data augmentations employed in many studies were performed on each of the three datasets such as random reflection, random rotation, random rescale, random translation along X-axis and random translation along Y-axis. After data augmentation, a total number of 10,000 COVID-19 positive CT san images were produced and 10,000 healthy individuals CT scan images were produced and a total number of 10,000 Common Pneumonia CT images were produced, in total, we generated 30,000 CT scan images for the study.

.

4.3.2. Training

In this study, seven pre-trained deep learning models AlexNet, GoogleNet, ResNet-50, ResNet-101, ShuffleNet, VGG16 and VGG19 which serves as feature extractor coupled with multi-class SVM classifier were used in classifying COVID-19, Common Pneumonia and Healthy individuals CT scan images, the pre-trained models serve as feature extractors while the multi-class SVM as classifier. The training is in two stages, firstly the training was carried out on the original set of CT scan images, while the second training was carried out on augmented CT scan images. The performance of the two training types was compared to find the best models. Figure 4.4. show the detailed training process in this study.

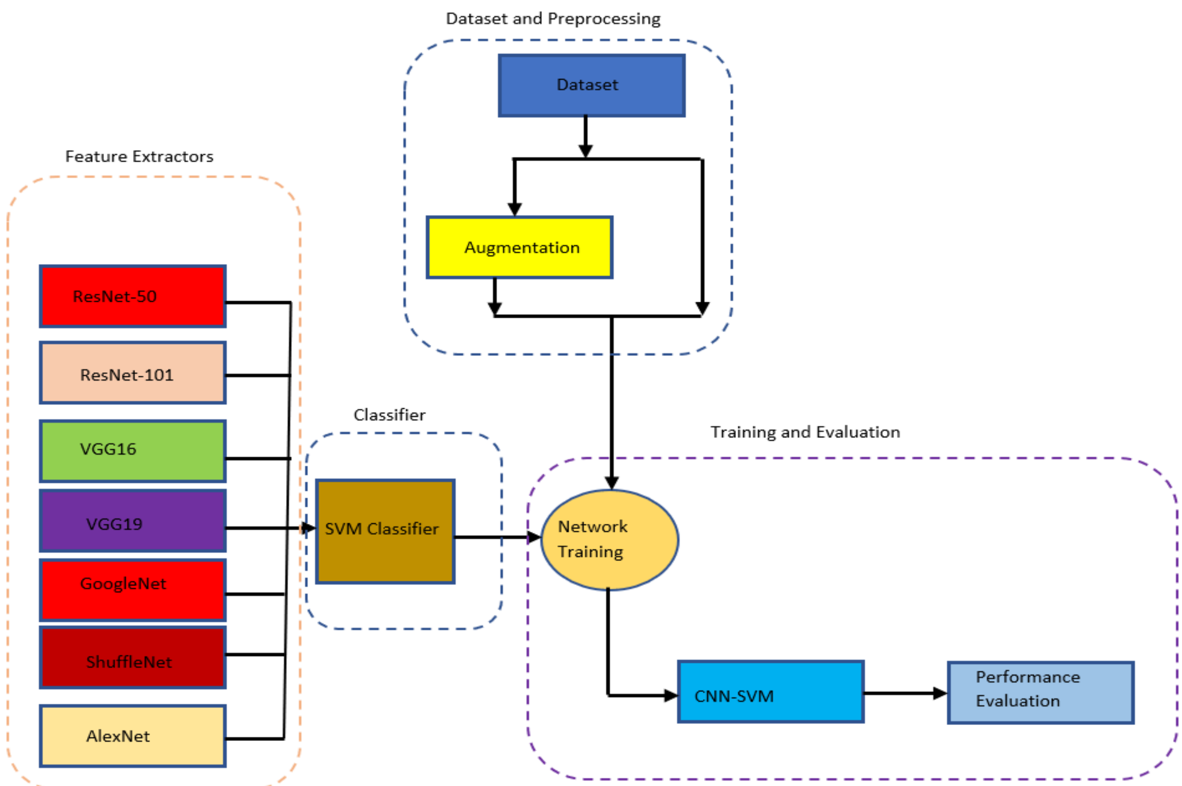


Figure 4.4: Training process

The images were resized to the size of 224 × 224 for GoogleNet, ResNet-50, ResNet-101, ShuffleNet, VGG16 and VGG19 and 227 × 227 for AlexNet, from the input of the pre-trained networks to the last pooling layer of the networks were used for feature extraction, the top of the pre-trained models which include the fully connected layers and the softmax classifier were replaced by multi-class SVM classifier, 805 of the total images were used for training and the 20% were used for testing.

4.4.Results and Discussion

In this study, seven pre-trained models were employed to classify COVID-19, Common Pneumonia and Healthy individuals CT scan images, the performance of the first and second training sets was compared to find the model with the best performance, also the performance of the best model between the two training was compared with the state of the art model that performed multi-class classification to detect COVID-19 on CT scan images.

The first training carried out on the original training set and the results of the model performance are presented in Table 4.2. The models with the best performance in terms of accuracy is the VGG16 with an accuracy of 93.8% followed by VGG19 and GoogleNet with an accuracy of 93.6% and 93.1% respectively, in terms of specificity, VGG16 achieves the highest with 0.943, also VGG16 achieves the highest AUC with 0.936, this show that VGG16 outperformed the remaining models trained in this model for the first training, the results are also visually presented in Figure 4.5 and Figure 4.6

Table 4.3: Proposed models performance for the first training

Models	Accuracy(%)	Sensitivity	Specificity	F1 Score	Precision	Yonden Index	AUC
Resnet-101	88.3	0.9	0.86	0.886	0.873	0.76	0.880
ResNet-50	91.9	0.918	0.918	0.918	0.919	0.836	0.918
GoogleNet	93.1	0.944	0.918	0.932	0.921	0.862	0.931
ShuffleNet	86	0.852	0.867	0.859	0.866	0.719	0.860

AlexNet	91	0.93	0.893	0.911	0.893	0.823	0.912
VGG16	93.8	0.934	0.943	0.939	0.944	0.877	0.939
VGG19	93.6	0.949	0.923	0.937	0.926	0.872	0.936

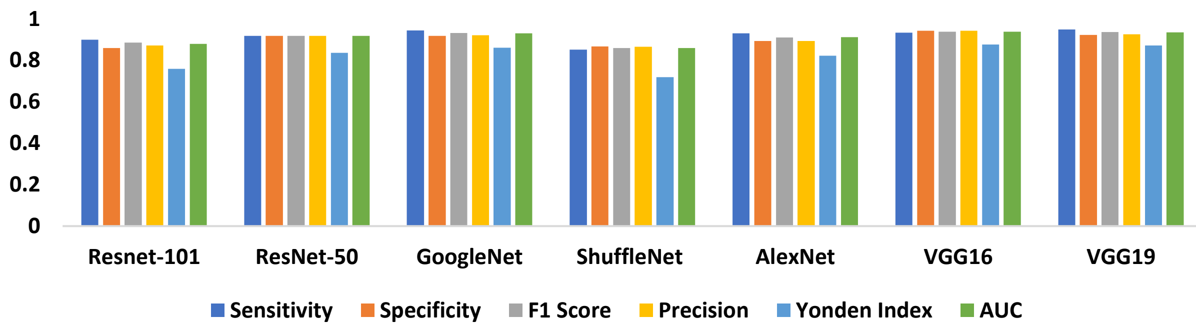


Figure 4.5: Models Performance for the first training

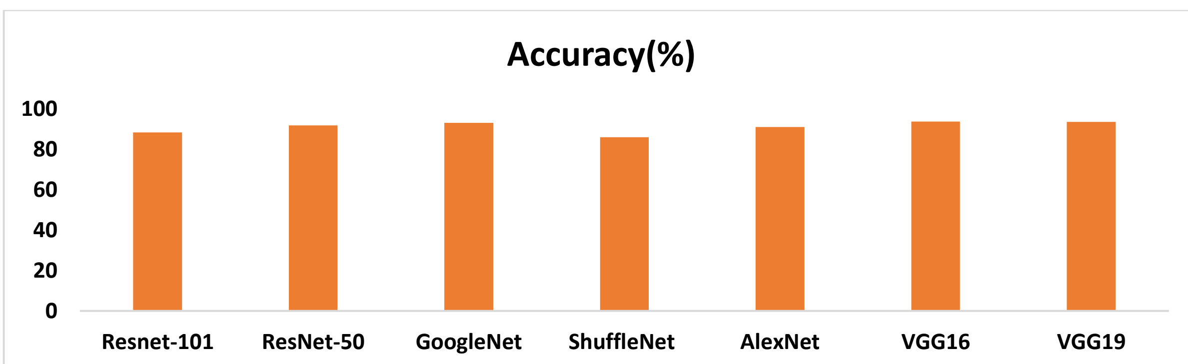


Figure 4.6: Models Accuracy for the first training

For the second training, the training was carried out on the augmented training set and the results are presented in Table 4.3. The models with the best performance in terms of accuracy is VGG19 with an accuracy of 96% followed by VGG16 and GoogleNet with an accuracy of 94.9 each. In terms of specificity and AUC, VGG19 outperformed the remaining models in the second training with 0.967 and 0.952 specificity and AUC respectively, the visual representation of the results is presented in Figure 4.7 and Figure 4.8

Table 4.3: Proposed models performance for the second training

Models	Accuracy(%)	Sensitivity	Specificity	F1 Score	Precision	Yonden Index	AUC
ResNet-101+Augmentation	93.4	0.912	0.948	0.905	0.898	0.86	0.930
ResNet-50+Augmentation	93	0.905	0.945	0.899	0.892	0.85	0.925
GoogleNet+Augmentation	94.9	0.926	0.964	0.927	0.928	0.89	0.945
AlexNet+Augmentation	92	0.886	0.941	0.885	0.883	0.827	0.914
ShuffleNet+Augmentation	87	0.846	0.892	0.848	0.851	0.738	0.869
VGG16+Augmentation	94.9	0.92	0.96	0.922	0.925	0.88	0.940
VGG19+Augmentation	96	0.936	0.967	0.935	0.934	0.903	0.952

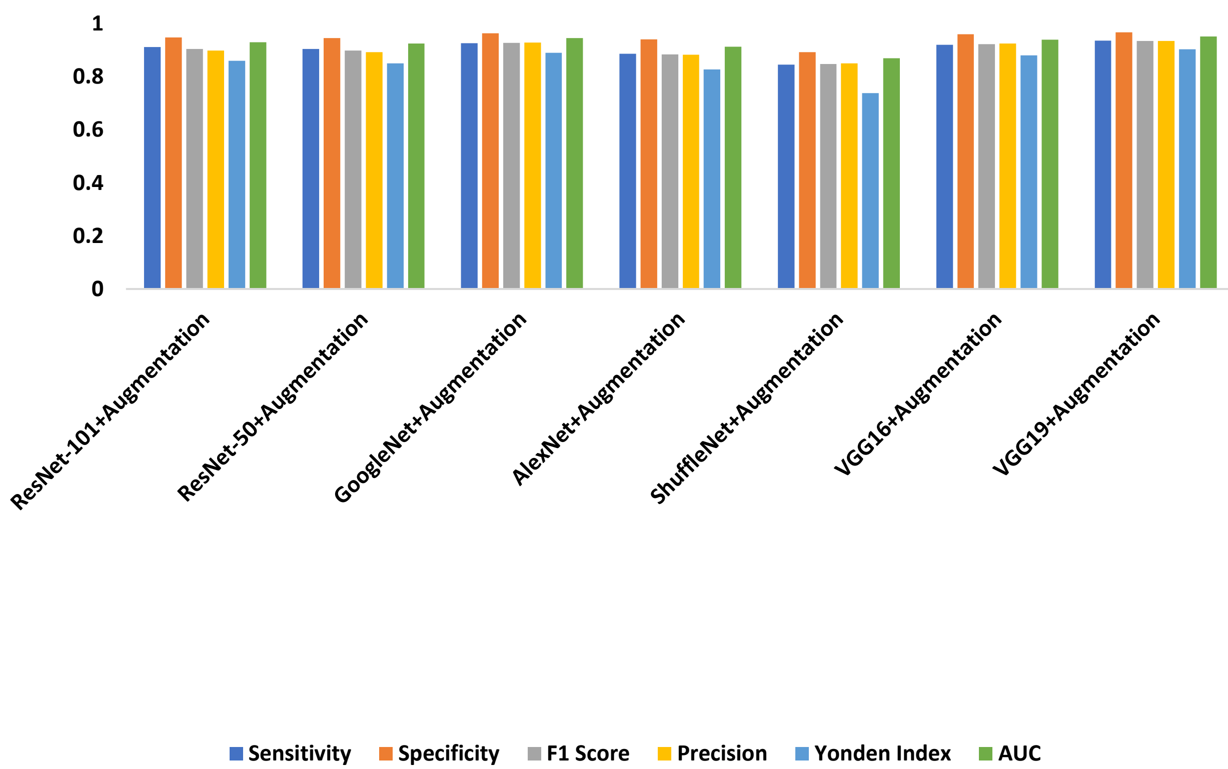


Figure 4.7: Models Performance for the second training

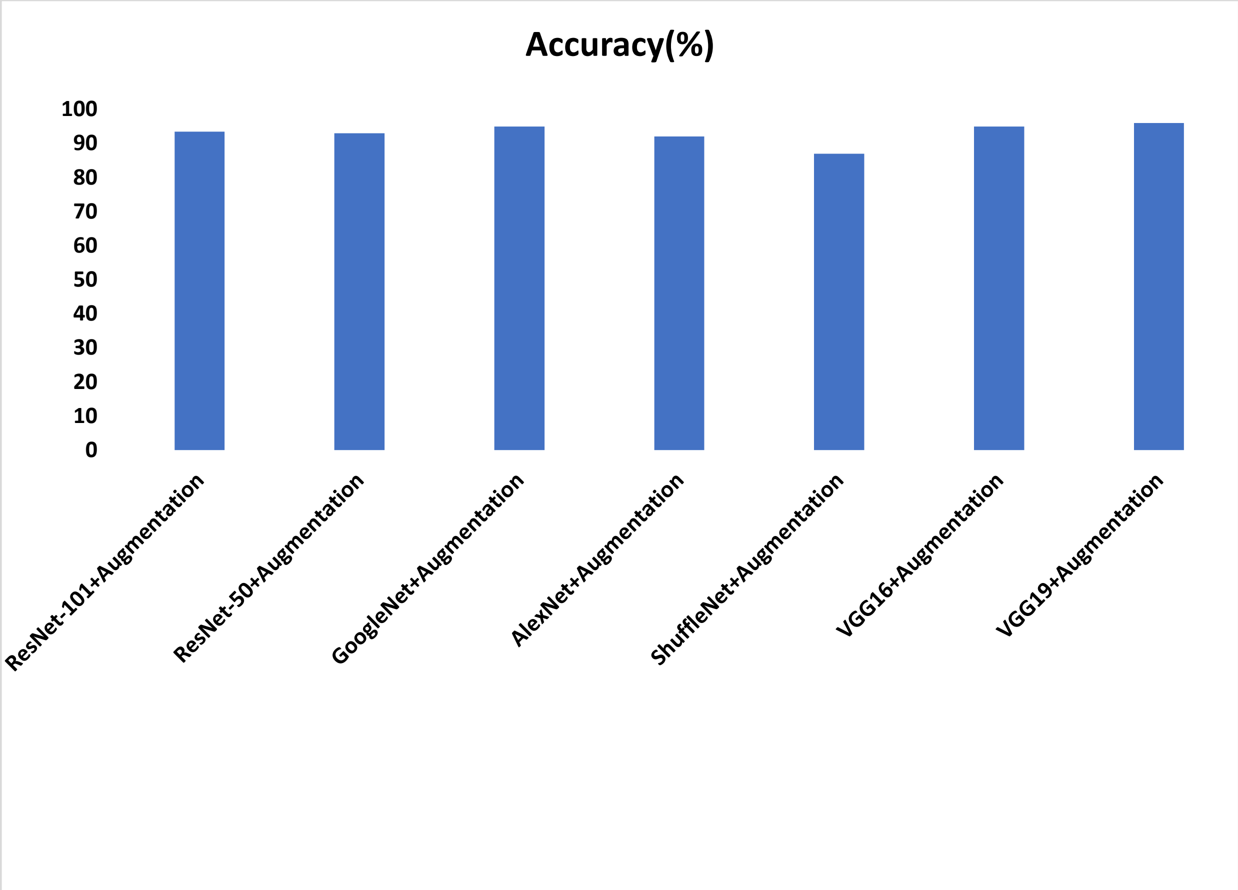


Figure 4.8: Models accuracy for the second training

The state of the art model (Amyar et al., 2020) proposed a new deep learning model in which they trained to classify CT scan images with three different classes COVID-19, other Pneumonia and Healthy patients, in training, they consider two different images resolutions, first training on 512 × 512 resolution CT scan images and the second training on 256 × 256 resolution CT scan images. The best models in (Amyar et al., 2020) achieve an accuracy of 94.67%, a sensitivity of 0.96, specificity of 0.92 and AUC of 0.97. in table 4.4. The proposed model's performance was compared to the state of the art model. The proposed architecture which utilizes pre-trained Deep leaning as feature extractor and SVM as classifier shows how great it can perform in detecting COVID-19 CT scan images.

Table 4.4: Proposed models performance with state of the art model

Ref.	Models	Accuracy(%)	Sensitivity	Specificity	F1 Score	Precision	Yonden Index	AUC
(Amyar et al.,	T1 & T2 & T3 512 × 512	91.13	0.94	0.85	NA	NA	NA	0.94
2020)	T1 & T2 & T3 256 x 256	94.67	0.96	0.92	NA	NA	NA	0.97
Proposed	Resnet-101	88.3	0.9	0.86	0.886	0.873	0.76	0.880
Models	ResNet-50	91.9	0.918	0.918	0.918	0.919	0.836	0.918
	GoogleNet	93.1	0.944	0.918	0.932	0.921	0.862	0.931
	ShuffleNet	86	0.852	0.867	0.859	0.866	0.719	0.860
	AlexNet	91	0.93	0.893	0.911	0.893	0.823	0.912
	VGG16	93.8	0.934	0.943	0.939	0.944	0.877	0.939
	VGG19	93.6	0.949	0.923	0.937	0.926	0.872	0.936
	ResNet-101+Augmentation	93.4	0.912	0.948	0.905	0.898	0.86	0.930
	ResNet-50+Augmentation	93	0.905	0.945	0.899	0.892	0.85	0.925
	GoogleNet+Augmentation	94.9	0.926	0.964	0.927	0.928	0.89	0.945
	AlexNet+Augmentation	92	0.886	0.941	0.885	0.883	0.827	0.914
	ShuffleNet+Augmentation	87	0.846	0.892	0.848	0.851	0.738	0.869
	VGG16+Augmentation	94.9	0.92	0.96	0.922	0.925	0.88	0.940
	VGG19+Augmentation	96	0.936	0.967	0.935	0.934	0.903	0.952

4.5.Performance Criterias

The performance criteria of a model describe how well the model performs in solving a problem. The performance of the two pre-trained models based on the binary and multi-class classification was compared using the frequently applied performance evaluation criteria, namely Validation Accuracy (ACC), Sensitivity(SN), Specificity(SP), F1 Score, Precision(PR), Yonden-Index and AUC. The evaluation criteria employed in this study are the widely adopted evaluation criteria employed in the classification of images(Aboughazala & Mohammed, 2020; Fan et al., 2020; M. M. Islam et al., 2021; Learning, 2020; Nour et al., 2020; Rahimzadeh & Attar, 2020b; Shaban et al., 2020; Shuai Wang, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao & Guo, Mengjiao Cai, Jingyi Yang, Yaodong Li, 2020; Shuai Wang et al., 2020).

4.5.1. Accuracy

Accuracy is a measure that gives an insight into how well the model learns and produces reliable results. It is the proportion of predictions that were provided correctly by the model. The accuracy of a model is the ratio of correctly predicted samples to the number of input samples. The number of correctly predicted samples is the sum of the number of true positives and false negatives

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{TP+FP+FN+TN}} \qquad (1)$$

4.5.2. Sensitivity or Recall

It is defined as the ability of a model to test correctly and identify patients with a disease as presented in Equation (2).

$$Recall/sensitivity\ = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (2)$$

4.5.3. Specificity

Specificity is a measure of how many negatives the trained model managed to capture out of the entire set of correctly predicted negative values by labelling the samples as negative. The relation for calculating specificity is presented in Equation 3.

$$Specificity\ = \frac{TN}{\mathrm{TN+FP}} \qquad (3)$$

4.5.4. F1-Score

F1-score is a measure of the balance between the precision and recall of a model. It is used to perform a statistical analysis of the test accuracy. The F1-score of a model lies between 0 and 1. It is said to be very good if its value lies near 1 and very bad if it is near 0. It is calculated by applying Equation 4

$F1\ Score = \frac{2\times(\text{Recall} \times \text{Precision})}{\text{Recall} + \text{Precision}}$ (4)

4.5.5. Yonden Index

Is the cut-point that optimizes the biomarker's distinguishing ability when equal weight is given to sensitivity and specificity. It also gives a summary of the receiver operating characteristic (ROC) curve, as presented in Equation 5.

$$Yonden\ Index = (\text{sensitivity} + \text{specificity}) - 1 \qquad (5)$$

4.5.6. Precision

Precision is a measure of how precise or accurate the model is in terms of positive classifications. In other words, it measures the number of true positives out of all the predicted positives. The relation for precision is presented in Equation 6

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \qquad (6)$$

True Positive (TP)

True Negative (TN)

False Positive (FP)

False Negative (FN)

CHAPTER 5
CONCLUSION

Detecting COVID-19 at an early stage in the time of the pandemic is very important, as we all know thoracic images are essential in detecting disease related to the lungs, CT scan images are unlike x-ray images which give only a latera view of the lungs, to generate CT scan images, more than 300 slices of different lungs level needs to be generated. This will be a tedious job for the radiologist at the time of the pandemic, a new method that will fasten and efficiently detect COVID-19 is needed.

Classifiers rely on how good features are extracted in order to classify medical images, that is why in this research we opted for the deep learning models and handcrafted features to classify CT scan images. Also, to improve the performance of the classifiers, a hybrid feature was proposed to improve the classification, the feature comprises deep learning and handcrafted features.

The limitations of the study are only two classifiers and three classes of CT scan images were considered. Also in the future, more classifiers and domain adaptation will be considered. In the future, IoT will also be considered in order to have more access to the trained model by many hospitals around the globe for easier detection of COVID-19.

REFERENCES

Abbasian, A., Rajabzadeh, A., & Acharya, U. R. (2020). *Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks*. *January*.

Aboughazala, L. M., & Mohammed, K. K. (2020). Automated Detection of Covid-19 Coronavirus Cases Using Deep Neural Networks with X-ray Images. *Al-Azhar University Journal of Virus Researches and Studies*, *2*(5), 1–12. https://doi.org/10.21608/aujv.2020.106722

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., & Xia, L. (2020). Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology*, *296*(2), E32–E40. https://doi.org/10.1148/radiol.2020200642

Alafif, T., Tehame, A. M., Bajaba, S., Barnawi, A., & Zia, S. (2021). Machine and deep learning towards covid-19 diagnosis and treatment: Survey, challenges, and future directions. *International Journal of Environmental Research and Public Health*, *18*(3), 1–24. https://doi.org/10.3390/ijerph18031117

Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P. B., Joe, B., & Cheng, X. (2020). Artificial intelligence and machine learning to fight covid-19. *Physiological Genomics*, *52*(4), 200–202. https://doi.org/10.1152/physiolgenomics.00029.2020

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *American Statistician*, *46*(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

Ammar, A., Mueller, P., Trabelsi, K., Chtourou, H., Boukhris, O., Masmoudi, L., Bouaziz, B., Brach, M., Schmicker, M., Bentlage, E., How, D., Ahmed, M., Aloui, A., Hammouda, O., Paineiras-Domingos, L. L., Braakmanjansen, A., Wrede, C., Bastoni, S., Pernambuco, C.

S., … Hoekelmann, A. (2020). Psychological consequences of COVID-19 home confinement: The ECLB-COVID19 multicenter study. *PLoS ONE*, *15*(11). https://doi.org/10.1371/journal.pone.0240204

Amyar, A., Modzelewski, R., Li, H., & Ruan, S. (2020). *Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation*. *January*.

Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, *43*(2), 635–640. https://doi.org/10.1007/s13246-020-00865-4

Butt, C., Gill, J., Chun, D., & Babu, B. A. (2020). Deep learning system to screen coronavirus disease 2019 pneumonia. *Applied Intelligence*, 1–29. https://doi.org/10.1007/s10489-020-01714-3

Chamola, V., Hassija, V., Gupta, V., & Guizani, M. (2020). A Comprehensive Review of the COVID-19 Pandemic and the Role of IoT, Drones, AI, Blockchain, and 5G in Managing its Impact. *IEEE Access*, *8*(April), 90225–90265. https://doi.org/10.1109/ACCESS.2020.2992341

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020a). Epidemiological and Clinical Characteristics of 99 Cases of 2019-Novel Coronavirus (2019-nCoV) Pneumonia in Wuhan, China. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3523861

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020b). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, *395*(10223), 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

Chen, X., Yao, L., Zhou, T., Dong, J., & Zhang, Y. (2021). Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images. *Pattern Recognition*, *113*. https://doi.org/10.1016/j.patcog.2021.107826

Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. Bin, Islam, K. R., Khan, M. S., Iqbal, A., Emadi, N. Al, Reaz, M. B. I., & Islam, M. T. (2020). Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access*, *8*, 132665–132676. https://doi.org/10.1109/ACCESS.2020.3010287

Chu, D. K. W., Pan, Y., Cheng, S. M. S., Hui, K. P. Y., Krishnan, P., Liu, Y., Ng, D. Y. M., Wan, C. K. C., Yang, P., Wang, Q., Peiris, M., & Poon, L. L. M. (2020). Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. *Clinical Chemistry*, *555*, 549–555. https://doi.org/10.1093/clinchem/hvaa029

Cnns, D. (2016). *Guest Editorial Deep Learning in Medical Imaging : Overview and Future Promise of an Exciting New Technique*. *35*(5), 1153–1159.

Cruz, J. A., & Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics*, *2*, 59–77. https://doi.org/10.1177/117693510600200030

Doyle, S., Hwang, M., Shah, K., Madabhushi, A., Feldman, M., & Tomaszeweski, J. (2007). Automated grading of prostate cancer using architectural and textural image features. *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*, 1284–1287. https://doi.org/10.1109/ISBI.2007.357094

Fan, D. P., Zhou, T., Ji, G. P., Zhou, Y., Chen, G., Fu, H., Shen, J., & Shao, L. (2020). Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Transactions on Medical Imaging*, *39*(8), 2626–2637. https://doi.org/10.1109/TMI.2020.2996645

Fehr, D., Veeraraghavan, H., Wibmer, A., Gondo, T., Matsumoto, K., Vargas, H. A., Sala, E., Hricak, H., & Deasy, J. O. (2015). Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proceedings of the National*

Academy of Sciences of the United States of America, *112*(46), E6265–E6273. https://doi.org/10.1073/pnas.1505935112

Gao, X. W., Hui, R., & Tian, Z. (2017). Classification of CT brain images based on deep learning networks. *Computer Methods and Programs in Biomedicine*, *138*, 49–56. https://doi.org/10.1016/j.cmpb.2016.10.007

Ghaderzadeh, M., & Asadi, F. (2021). Deep Learning in the Detection and Diagnosis of COVID-19 Using Radiology Modalities: A Systematic Review. *Journal of Healthcare Engineering*, *2021*. https://doi.org/10.1155/2021/6677314

Ghassemi, N., Shoeibi, A., & Rouhani, M. (2020). Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomedical Signal Processing and Control*, *57*, 101678. https://doi.org/10.1016/j.bspc.2019.101678

Ghoshal, B., & Tucker, A. (2020). *Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection*. 1–14. http://arxiv.org/abs/2003.10769

Haque, A. B., & Rahman, M. (2020). *Augmented COVID-19 X-ray Images Dataset (Mendely) Analysis using Convolutional Neural Network and Transfer Learning*. *19*(April). https://doi.org/10.13140/RG.2.2.20474.24003

He, K., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. https://doi.org/10.1109/CVPR.2016.90

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 2261–2269.

https://doi.org/10.1109/CVPR.2017.243

Hussein, S., Kandel, P., Bolan, C. W., Wallace, M. B., & Bagci, U. (2019). Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches. *IEEE Transactions on Medical Imaging*, *38*(8), 1777–1787. https://doi.org/10.1109/TMI.2019.2894349

Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, M. S., Ali, K., John, C. N., Hussain, M. I., & Nabeel, M. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked*, *20*, 100378. https://doi.org/10.1016/j.imu.2020.100378

Islam, M. M., Karray, F., Alhajj, R., & Zeng, J. (2021). A Review on Deep Learning Techniques for the Diagnosis of Novel Coronavirus (COVID-19). *IEEE Access*, *9*, 30551–30572. https://doi.org/10.1109/ACCESS.2021.3058537

Islam, M. Z., Islam, M. M., & Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*, *20*, 100412. https://doi.org/10.1016/j.imu.2020.100412

Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., Pasupuleti, S., Mishra, R., Pillai, S., & Jo, O. (2020). COVID-19 patient health prediction using boosted random forest algorithm. *Frontiers in Public Health*, *8*(July), 1–9. https://doi.org/10.3389/fpubh.2020.00357

Jain, R., Gupta, M., Taneja, S., & Hemanth, D. J. (2021). Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Applied Intelligence*, *51*(3), 1690–1700. https://doi.org/10.1007/s10489-020-01902-1

KrizhevskyAlex, Ilya Sutskever, G. E. H. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In T. F. Gonzalez (Ed.), *Handbook of Approximation Algorithms and Metaheuristics*. Chapman and Hall/CRC. https://doi.org/10.1201/9781420010749

Kroft, L. J. M., Van Der Velden, L., Girón, I. H., Roelofs, J. J. H., De Roos, A., & Geleijns, J. (2019). Added value of ultra-low-dose computed tomography, dose equivalent to chest x-ray radiography, for diagnosing chest pathology. *Journal of Thoracic Imaging*, *34*(3), 179–186. https://doi.org/10.1097/RTI.0000000000000404

Learning, T. (2020). *SS symmetry Within the Lack of Chest COVID-19 X-ray Dataset : A Novel Detection Model Based on GAN and Deep*.

Lee, C. Y., Xie, S., Gallagher, P. W., Zhang, Z., & Tu, Z. (2015). Deeply-supervised nets. *Journal of Machine Learning Research*, *38*, 562–570.

Li, Y., Zhang, H., & Shen, Q. (2017). Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, *9*(1). https://doi.org/10.3390/rs9010067

Liu, H., Liu, F., Li, J., Zhang, T., Wang, D., & Lan, W. (2020). Clinical and CT imaging features of the COVID-19 pneumonia: Focus on pregnant women and children. *Journal of Infection*, *80*(5), e7–e13. https://doi.org/10.1016/j.jinf.2020.03.007

Liu, R., Han, H., Liu, F., Lv, Z., Wu, K., Liu, Y., Feng, Y., & Zhu, C. (2020). Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clinica Chimica Acta*, *505*(March), 172–175. https://doi.org/10.1016/j.cca.2020.03.009

Loey, M., Smarandache, F., & Khalifa, N. E. M. (2020). *A Deep Transfer Learning Model with Classical Data Augmentation and CGAN to Detect COVID-19 from Chest CT Radiography Digital Images*. *April*, 1–17. https://doi.org/10.20944/preprints202004.0252.v1

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., … Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, *395*(10224), 565–574.

https://doi.org/10.1016/S0140-6736(20)30251-8

Luz, E., Silva, P., Silva, R., Silva, L., Guimarães, J., Miozzo, G., Moreira, G., & Menotti, D. (2021). Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. *Research on Biomedical Engineering*, *July*. https://doi.org/10.1007/s42600-021-00151-6

Maghded, H. S., Ghafoor, K. Z., Sadiq, A. S., Curran, K., Rawat, D. B., & Rabie, K. (2020). A Novel AI-enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study. *Proceedings - 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science, IRI 2020*, 180–187. https://doi.org/10.1109/IRI49571.2020.00033

Mahmud, T., Rahman, M. A., & Fattah, S. A. (2020). CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Computers in Biology and Medicine*, *122*(May), 103869. https://doi.org/10.1016/j.compbiomed.2020.103869

Mohammed, M. N., Hazairin, N. A., Syamsudin, H., Al-Zubaidi, S., Sairah, A. K., Mustapha, S., & Yusuf, E. (2020). 2019 novel coronavirus disease (Covid-19): Detection and diagnosis system using iot based smart glasses. *International Journal of Advanced Science and Technology*, *29*(7 Special Issue), 954–960.

Mohammed, M. N., Syamsudin, H., Al-Zubaidi, S., Sairah, A. K., Ramli, R., & Yusuf, E. (2020). Novel covid-19 detection and diagnosis system using iot based smart helmet. *International Journal of Psychosocial Rehabilitation*, *24*(7), 2296–2303. https://doi.org/10.37200/IJPR/V24I7/PR270221

Multiple, A. D. (2020). *Accurate Screening of COVID-19 Using Attention-Based Deep 3D Multiple Instance Learning*. *39*(8), 2584–2594.

Murphy, M. P. A. (2020). COVID-19 and emergency eLearning: Consequences of the securitization of higher education for post-pandemic pedagogy. *Contemporary Security Policy*, *41*(3), 492–505. https://doi.org/10.1080/13523260.2020.1761749

Neher, R. A., Dyrdak, R., Druelle, V., Hodcroft, E. B., & Albert, J. (2020). Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Medical Weekly*, *150*, w20224. https://doi.org/10.4414/smw.2020.20224

Nguyen, T. T., Nguyen, Q. V. H., Nguyen, D. T., Hsu, E. B., Yang, S., & Eklund, P. (2020). *Artificial Intelligence in the Battle against Coronavirus (COVID-19): A Survey and Future Research Directions*. 1–14. http://arxiv.org/abs/2008.07343

Niu, X. X., & Suen, C. Y. (2012). A novel hybrid CNN-SVM classifier for recognizing handwritten digits. *Pattern Recognition*, *45*(4), 1318–1325. https://doi.org/10.1016/j.patcog.2011.09.021

Nour, M., Cömert, Z., & Polat, K. (2020). A Novel Medical Diagnosis model for COVID-19 infection detection based on Deep Features and Bayesian Optimization. *Applied Soft Computing Journal*, *97*, 106580. https://doi.org/10.1016/j.asoc.2020.106580

Oakden-Rayner, L., Carneiro, G., Bessen, T., Nascimento, J. C., Bradley, A. P., & Palmer, L. J. (2017). Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Scientific Reports*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-01931-w

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1140–1152. https://doi.org/10.1016/j.neubiorev.2012.01.004

Oruc, B. E., Baxter, A., Keskinocak, P., Asplund, J., & Serban, N. (2021). Homebound by COVID19: the benefits and consequences of non-pharmaceutical intervention strategies. *BMC Public Health*, *21*(1), 1–16. https://doi.org/10.1186/s12889-021-10725-9

Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Rajendra Acharya, U. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine*, *121*(April), 103792. https://doi.org/10.1016/j.compbiomed.2020.103792

Parmar, N. S., Yim, H., & Choi, J. W. (2017). Critical increase in Na-doping facilitates acceptor band movements that yields ∼180 meV shallow hole conduction in ZnO bulk crystals. *Scientific Reports*, *7*(November 2016), 1–7. https://doi.org/10.1038/srep44196

Pinter, G., Felde, I., Mosavi, A., Ghamisi, P., & Gloaguen, R. (2020). COVID-19 Pandemic Prediction for Hungary; A Hybrid Machine Learning Approach. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3590821

Pirouz, B., Haghshenas, S. S., Haghshenas, S. S., & Piro, P. (2020). Investigating a serious challenge in the sustainable development process: Analysis of confirmed cases of COVID-19 (new type of Coronavirus) through a binary classification using artificial intelligence and regression analysis. *Sustainability (United States)*, *12*(6). https://doi.org/10.3390/su12062427

Pratap, R., Javaid, M., Haleem, A., & Suman, R. (2020). *Internet of things (IoT) applications to fight against COVID-19 pandemic*. *January*. https://doi.org/10.1016/j.dsx.2020.04.041

Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, *32*(NeurIPS).

Rahimzadeh, M., & Attar, A. (2020a). A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informatics in Medicine Unlocked*, *19*, 100360. https://doi.org/10.1016/j.imu.2020.100360

Rahimzadeh, M., & Attar, A. (2020b). *A New Modified Deep Convolutional Neural Network for Detecting COVID-19 from X-ray Images*. https://doi.org/10.1016/j.imu.2020.100360

Roth, H. R., Lu, L., Member, S., Liu, J., Yao, J., Seff, A., Cherry, K., Kim, L., & Summers, R. M. (2015). *Improving Computer-aided Detection using Convolutional Neural Networks and Random View Aggregation*. 1–12. https://doi.org/10.1109/TMI.2015.2482920

Rustam, F., Reshi, A. A., Mehmood, A., Ullah, S., On, B. W., Aslam, W., & Choi, G. S. (2020). COVID-19 Future Forecasting Using Supervised Machine Learning Models. *IEEE Access*, *8*, 101489–101499. https://doi.org/10.1109/ACCESS.2020.2997311

Sa, Z., Ozsoz, M., & Saleh, A. (2021). C-SVR Crispr : Prediction of CRISPR / Cas12 guideRNA activity using deep learning models. *Alexandria Engineering Journal*, *60*(4), 3501–3508. https://doi.org/10.1016/j.aej.2021.02.007

Shaban, W. M., Rabie, A. H., Saleh, A. I., & Abo-Elsoud, M. A. (2020). A new COVID-19 Patients Detection Strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier. *Knowledge-Based Systems*, *205*. https://doi.org/10.1016/j.knosys.2020.106270

Shuai Wang, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao, J., & Guo, Mengjiao Cai, Jingyi Yang, Yaodong Li, X. M. (2020). *A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)*. 1–27.

Simon, P., Uma, V., Simon, P., & Uma, V. (2020). Deep Learning based Feature Extraction for Texture Classification Deep Learning based Feature Extraction for Texture Classification. *Procedia Computer Science*, *171*(2019), 1680–1687. https://doi.org/10.1016/j.procs.2020.04.180

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.

Singh, D., Kumar, V., Vaishali, & Kaur, M. (2020). Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. *European Journal of Clinical Microbiology and Infectious Diseases*, *39*(7), 1379–1389. https://doi.org/10.1007/s10096-020-03901-z

Soares, E., Angelov, P., Biaso, S., Higa Froes, M., & Kanda Abe, D. (2020). *SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification*. 1–8. https://doi.org/10.1101/2020.04.24.20078584

Strunk, J. L., Temesgen, H., Andersen, H., & Packalen, P. (2014). *Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review Authors: 80*(2), 1–8. https://doi.org/10.14358/PERS.80.2.000

Szegedy, C., Reed, S., Sermanet, P., Vanhoucke, V., & Rabinovich, A. (n.d.). *Going deeper with convolutions*. 1–12.

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., Liang, J., & Member, S. (2016). *Convolutional Neural Networks for Medical Image Analysis : Full Training or Fine Tuning ? 35*(5), 1299–1312.

Tao Yan, Pak Kin Wong, Hao Ren c, Huaqiao Wang, Jiangtao Wang, Y. L. (n.d.). *COVID-19 and common pneumonia chest CT dataset*. https://doi.org/10.17632/3y55vgckg6.1

Toğaçar, M., Ergen, B., & Cömert, Z. (2020). COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Computers in Biology and Medicine*, *121*(May). https://doi.org/10.1016/j.compbiomed.2020.103805

Tuli, S., Tuli, S., Tuli, R., & Gill, S. S. (2020). Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. *Internet of Things*, *11*(May), 100222. https://doi.org/10.1016/j.iot.2020.100222

Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, *14*(4), 337–339. https://doi.org/10.1016/j.dsx.2020.04.012

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-turjman, F., & Pinheiro, P. R. (2020). *CovidGAN : Data Augmentation using Auxiliary Classifier GAN for Improved Covid-19*

Detection. 1–9. https://doi.org/10.1109/ACCESS.2020.2994762

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., & Peng, Z. (2020). Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - Journal of the American Medical Association*, *323*(11), 1061–1069. https://doi.org/10.1001/jama.2020.1585

Wang, J., Ding, H., Bidgoli, F. A., Zhou, B., Iribarren, C., Molloi, S., & Baldi, P. (2017). Mammograms With Deep Learning. *IEEE Transactions on Medical Imaging*, *36*(5), 1172–1181.

Wang, L., Lin, Z. Q., & Wong, A. (2020). COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-76550-z

Wang, Shuai, Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., & Xu, B. (2020). *A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)*. 1–19. https://doi.org/10.1101/2020.02.14.20023028

Wang, Shuihua, Sun, J., Chen, Y., Zhang, Y., & Mehmood, I. (2019). *Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling*. *May 2018*, 1–16. https://doi.org/10.1002/cpe.5130

Wang, Shuihua, Tang, C., Sun, J., Zhang, Y., Kingdom, U., & Kingdom, U. (2019). *Cerebral micro-bleeding detection based on Densely connected neural network*. https://doi.org/10.3389/fnins.2019.00422

Wang, W. C., Xu, D. M., Chau, K. W., & Chen, S. (2013). Improved annual rainfall-runoff forecasting using PSO-SVM model based on EEMD. *Journal of Hydroinformatics*, *15*(4), 1377–1390. https://doi.org/10.2166/hydro.2013.134

Weng, W., & Zhu, X. (2021). INet: Convolutional Networks for Biomedical Image

Segmentation. *IEEE Access*, *9*, 16591–16603. https://doi.org/10.1109/ACCESS.2021.3053408

Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, *45*(2), 107–114. https://doi.org/10.1080/17439884.2020.1761641

Xie, M., & Chen, Q. (2020). Insight into 2019 novel coronavirus — An updated interim review and lessons from SARS-CoV and MERS-CoV. *International Journal of Infectious Diseases*, *94*, 119–124. https://doi.org/10.1016/j.ijid.2020.03.071

Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., Shi, Q., Jin, S., Zhang, L., & You, Z. (2020). *COVID-19 Chest CT Image Segmentation -- A Deep Convolutional Neural Network Solution*. 1–10. http://arxiv.org/abs/2004.10987

Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., & Xie, P. (n.d.). *COVID-CT-Dataset : A CT Image Dataset about COVID-19*. 1–14.

Ye, Z., Zhang, Y., Wang, Y., Huang, Z., & Song, B. (2020). Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *European Radiology*, *30*(8), 4381–4389. https://doi.org/10.1007/s00330-020-06801-0

Zhai, P., Ding, Y., Wu, X., Long, J., Zhong, Y., & Li, Y. (2020). The epidemiology, diagnosis and treatment of COVID-19. *International Journal of Antimicrobial Agents*, *55*(5). https://doi.org/10.1016/j.ijantimicag.2020.105955

Zhang, Q., Xiao, Y., Dai, W., Suo, J., Wang, C., Shi, J., & Zheng, H. (2016). Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics*, *72*, 150–157. https://doi.org/10.1016/j.ultras.2016.08.004

Zhang, X., Lin, M., & Sun, J. (n.d.). *ShuffleNet : An Extremely Efficient Convolutional Neural Network for Mobile Devices*. 6848–6856.

Zheng, N., Du, S., Wang, J., Zhang, H., Cui, W., Kang, Z., Yang, T., Lou, B., Chi, Y., Long, H., Ma, M., Yuan, Q., Zhang, S., Zhang, D., Ye, F., & Xin, J. (2020). Predicting COVID-19 in China Using Hybrid AI Model. *IEEE Transactions on Cybernetics*, *50*(7), 2891–2904. https://doi.org/10.1109/TCYB.2020.2990162

Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., & Summers, R. M. (2021). A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises. *Proceedings of the IEEE*, *109*(5), 820–838. https://doi.org/10.1109/JPROC.2021.3054390

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017

Zisserman Karen, S. A. (2015). *VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION*. 1–14.

APPENDIX 1

ETHICAL APPROVAL LETTER

TO INSTITUTE OF GRADUATE STUDIES

REFERENCE: AUWALU SALEH MUBARAK (20178451)

I would like to inform you that the above candidate is one of our postgraduate students in Electrical and Electronics Engineering department he is taking thesis under my supervision and the thesis entailed: **EFFICIENT COVID-19 DETECTION ON COMPUTED TOMOGRAPHY IMAGES.** The data used in his thesis does not require any ethical report.

Please do not hesitate to contact me if you have any further queries or questions.

Thank you very much indeed.

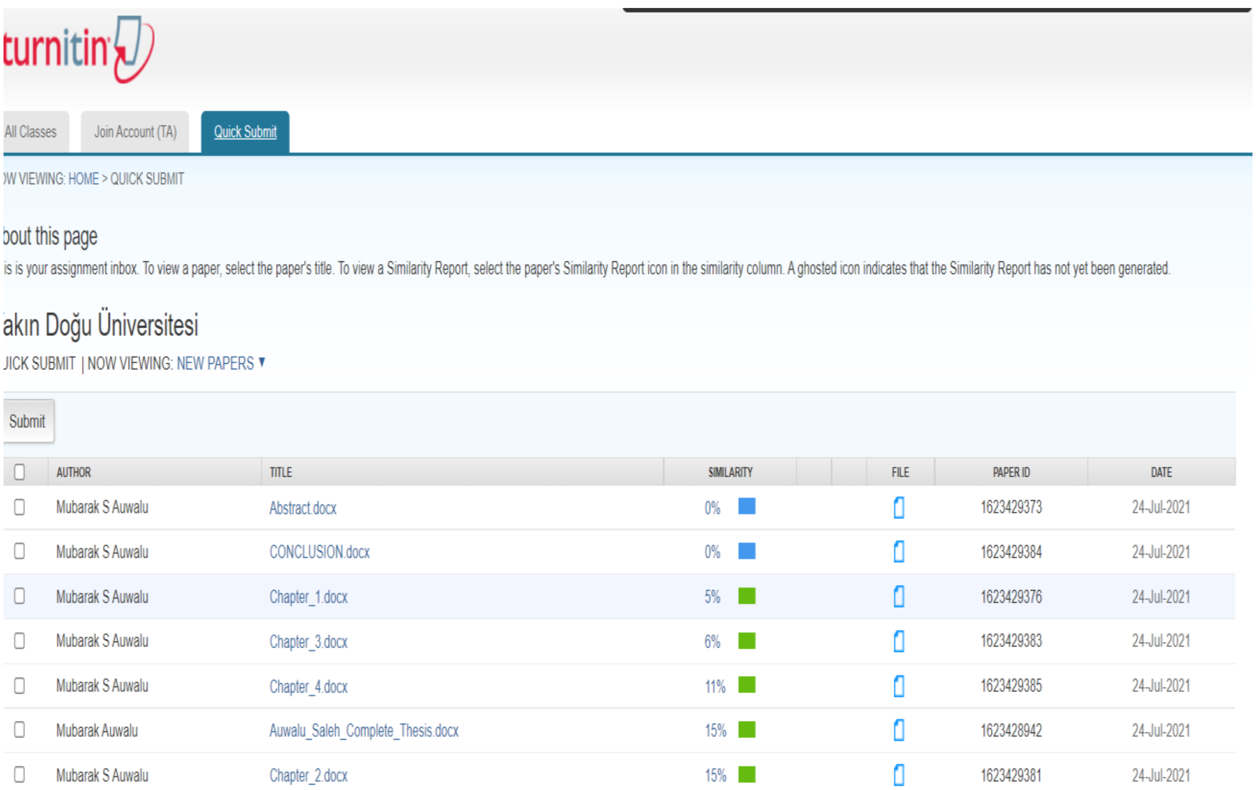
Best Regards,

Assoc. Prof. Dr. Sertan Serte

Near East University,
Electrical and Electronics Engineering Department,
Near East Boulevard, ZIP: 99138
Nicosia / TRNC, North Cyprus,
Mersin 10 – Turkey.
Email: sertan.serte@neu.edu.tr

APPENDIX 2

SIMILARITY REPORT


turnitin

All Classes | Join Account (TA) | Quick Submit

)W VIEWING: HOME > QUICK SUBMIT

bout this page

is is your assignment inbox. To view a paper, select the paper's title. To view a Similarity Report, select the paper's Similarity Report icon in the similarity column. A ghosted icon indicates that the Similarity Report has not yet been generated.

akın Doğu Üniversitesi

JICK SUBMIT | NOW VIEWING: NEW PAPERS ▼

Submit

AUTHOR	TITLE	SIMILARITY	FILE	PAPER ID	DATE
Mubarak S Auwalu	Abstract.docx	0%		1623429373	24-Jul-2021
Mubarak S Auwalu	CONCLUSION.docx	0%		1623429384	24-Jul-2021
Mubarak S Auwalu	Chapter_1.docx	5%		1623429376	24-Jul-2021
Mubarak S Auwalu	Chapter_3.docx	6%		1623429383	24-Jul-2021
Mubarak S Auwalu	Chapter_4.docx	11%		1623429385	24-Jul-2021
Mubarak Auwalu	Auwalu_Saleh_Complete_Thesis.docx	15%		1623428942	24-Jul-2021
Mubarak S Auwalu	Chapter_2.docx	15%		1623429381	24-Jul-2021

APPENDIX 3

CURRICULUM VITAE

CURRICULUM VITAE

Auwalu Saleh Mubarak In 2012, he received a bachelor's degree in Electrical Engineering with a Specialization in Electronics and Control from Cape Breton University, and in 2015, he received a master's degree in Electrical and Electronics Engineering with a specialization in Instrumentation and Control from Sharda University. Mubarak Auwalu is a PhD student at Near East University studying electrical engineering. Artificial Intelligence, BlockChain, and the Internet of Things are among his research interests. Mubarak Auwalu is the author of more than ten articles in prestigious journals and book series. In the year 2021, he has reviewed seven manuscripts for reputable journals like Expert Systems, Distributed and Parallel Databases, and Mobile Networks and Application.

PERSONAL DATA

Name:	Mubarak Auwalu Saleh
Date of Birth:	28^{th} August,1989
Nationality:	Nigerian
Place of Birth:	Kano
State of Origin:	Kano
Permanent Address:	YA 275 Zoo Road Housing Estate Kano.
Current Address:	Block 16, Room 1614, NEU Cyprus
Mobile Number:	05338717889
Marital Status:	Married

Emails: mubarakauwal@gmail.com, auwalusaleh.mubarak@neu.edu.tr

Skype ID: mubarak.auwal

QUALIFICATIONS

INSTITUTIONS	QUALIFICATIONS	DATES
Sharda University	M.Tech Electrical and Electronics engineering(Instrumentation and Control)	2014-2015
Cape Breton University	B.ENG Electrical Engineering. (Electronics and Control)	2008-2012

OTHER QUALIFICATIONS

- Three-month Industrial automation training(PLC,SCADA,HMI,VFD) at CETPA info-Tech India (2015)
- Wireless and telecom training at CETPA info-tech India (2015)
- Autocad certificate at Inter System Computer Training School (2006)
- Basic computer appreciation for JSS/SSS students at Hands-on Institute Information Technology (2002)
- Advance computer appreciation for JSS/SSS students at Hands-on Institute Information Technology (2004)

WORKING EXPERIENCE

- Near East University as a Research Assistant
- Kano University of Science and Technology, Kano. As Lecturer to date.
- Kano State Project Monitoring and Evaluation Directorate as an Electrical Engineer. (2013- 2016).
- National Youth Service Corps primary assignment at Ministry of Works, Housing and Transportation as an Electrical Engineer in Street and Traffic light department. (2012-2013)

Awards

- Kano State masters scholarship 2014
- Kano State PhD. Scholarship 2018

PUBLICATIONS

- A. S. Mubarak, Z. Sa'id Ameen, P. Tonga, and F. Al-Turjman, "Smart Tourism: A Proof of Concept For Cyprus Museum of Modern Arts In The IoT Era," pp. 49–53, 2021, doi: 10.1109/icaiot53762.2021.00016.
- A. S. Mubarak and S. Serte, "Local binary pattern and deep learning feature extraction fusion for COVID-19 detection on computed tomography images," no. September, pp. 1–13, 2021, doi: 10.1111/exsy.12842.
- Mubarak Auwalu Saleh, Sertan Serte, Fadi Al-Turjman, R.A. Abdulkadir, Zubaida Sa'id Ameen, Mehmet ozsoz. Deep learning-based feature extraction coupled with multi-class SVM for COVID-19 detection in the IoT eraInternational Journal of Nano Technology(2021). DOI: 10.1504/IJNT.2021.10040115
- Mubarak, A. S., Esmaili, P., Ameen, Z. S., Abdulkadir, R. A., Gaya, M. S., Ozsoz, M., Saini, G., & Abba, S. I. (2021). Metro-environmental data approach for the prediction of chemical oxygen demand in new nicosia wastewater treatment plant. *Desalination and Water Treatment*, *221*(May), 31–40. https://doi.org/10.5004/dwt.2021.27049
- C-SVR Crispr: Prediction of CRISPR/Cas12 guideRNA activity using deep learning models, Zubaida Sa'id Ameen, Mehmet Ozsoz, Auwalu Saleh Mubarak, Fadi Al Turjman, SertanSerte, Alexandria Engineering Journal, 2020, DOI: 10.1016/j.aej.2021.02.007
- Estimation of water quality index using artificial intelligence approaches and multi-linear regression, Muhammad Sani Gaya, Sani Isah Abba, Aliyu Muhammad Abdu, Abubakar Ibrahim Tukur, Mubarak Auwal Saleh, Parvaneh Esmaili, Norhaliza Abdul Wahab, IAES International Journal of Artificial Intelligence, Vol. 9, No. 1, March 2020, pp. 126~134 ISSN: 2252-8938, DOI: 10.11591/ijai.v9.i1.pp126-134
- Development of CNN Model for Prediction of CRISPR/Cas12 Guide RNA Activity ZS Ameen, AS Mubarak, A Süleyman, O Mehmet International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions, 697-703, Springer, Cham
- Neuro-fuzzy ensemble techniques for the prediction of turbidity in water treatment plant, SI Abba, RA Abdulkadir, MS Gaya, MA Saleh, P Esmaili, MB Jibril 2019 2nd International Conference of the IEEE Nigeria Computer Chapter 2019

- Modelling of Uncertain System: A comparison study of Linear and Non-Linear Approaches, SI Abba, MS Gaya, ML Yakubu, MU Zango, RA Abdulkadir, MA Saleh, AN Hamza, Ukashatu Abubakar, AI Tukur, NA Wahab 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), May 2019,DOI: 10.1109/I2CACIS.2019.8825085
- Application of data driven algorithms for the forecasting of non-linear parameter, S.I. Abba, A.S. Maihula, M.B. Jibril, A.M. Sunusi, M.A. Ahmad, M. A. Saleh, International Journal of Recent Engineering Science (IJRES), Volume 6 Issue 2, March-april, 2019
- Implementation of Microcontroller Based Distance Relay Abubakar Isa, H. K. Verma, Abdurrahman Shuaibu, and Mubarak Auwal Saleh, EJECE, European Journal of Electrical and Computer Engineering Vol. 2, No. 5, July 2018, DOI: http://dx.doi.org/10.24018/ejece.2018.2.5.31
- Mubarak Auwwal Saleh, Design and construction of RFID Car Security System, ISSN(online)2456-7361, IJSES, Volume1, Issue 6,2017
- Auwalu Saleh Mubarak "Development of Automation System for Residential Complex for the Elderly" International Journal of Electrical and Electronics Engineering. ISSN 2321-2055, Volume 07 june (2015)
- A.S. Mubarak, A.S Hassan, N.H. Umar, and M. Nasiru " An Analytical Study of Power System Under the Fault Conditions Using different Methods of Fault Analysis" Advance Research in Electrical and Electronic Engineering, Print ISSN: 2349-5804; Online ISSN; 2349-5812 Volume 2, Number 10 April-June (2015)pp.113-119
- Mubarak A. "comparative assessment of reliability of zinox laptop computers in different range of years" International journal of Advance Research in Science and Engineering ISSN: 2319-8354, Volume 04, Issue S.I (01) September 2015.
- Saleh M.A. Design and Simulation of Circular patch Antenna with a Reconfigurable Polarization. Bayero Journal of Engineering and Technology(BJET), ISSN2449-0539, Vol12 no.2 August 2017

MEMBERSHIP

- Member Council for the Regulation of Engineering in Nigeria (COREN:R34313)
- IEEE Member (93854889)
- International Association of Engineers (160015)

CONFERENCES AND WORKSHOPS ATTENDED

- International Conference on Innovative Research in "Mechanical, Electrical, Electronics, Civil, Computer Science and Information Technology (MECIT-2015)

- International Conference on Computing Communication and Automation (ICCCA 2015), (IEEE Conference Record No 36179)
- National Conference on Emerging Trends of Electronics and Communication Engineering (NCECE)2015
- 2 days workshop on robotics by GIZMO AND CETPA India
- National conference on emerging trends of electronics and communication engineering (2015) India
- One day workshop on panasonic (PLC,HMI,SCADA) India
- One day workshop on Embedded System by Ducat, India.
- One day workshop on Industrial Automation by Ducat, India.

REFEREES

- Name: Dr. Mahadi Lawan Yakub

  Address: Faculty of Engineering Kano University of Science and Technology

  Mobile number:08036782425

- Name: Dr. Dahiru Sani

  Address: Faculty of Technology Bayero University Kano.

  Mobile number: 08099045666