Nat. Volatiles & Essent. Oils, 2021; 8(4): 5213-5236

Prediction Of Covid-19 Cases At The Early Stageusing Deep Learning Method Taken Over CT Scan Images

Prashanth Kambli[1], Nikita Juana Santiago[2*], Naresh.E[3]

[1]Assistant Professor, Department of Information Science & Engineering, M S Ramaiah Institute of Technology, Bangalore-560054, Karnataka, India

[2*]PG student, Department of Information Science & Engineering, M S Ramaiah Institute of Technology, Bangalore-560054, Karnataka, India

[3]Assistant Professor, Department of Information Science & Engineering, M S Ramaiah Institute of Technology, Bangalore-560054, Karnataka, India

ABSTRACT

A highly transmissible disease called Corona Virus Disease (COVID-19) causeddue to a syndrome found in respiratory action that is Severe Acute Coronavirus 2(SARS-CoV-2). People of old age, low immunity, medical issues and history of respiratory problems mainly related to lungsbecome more exposed to COVID-19. The best way to save a person who is suffering from Covid-19 is early prediction of the virus and early treatment. The normal testing methods for Covid-19 takes time to produce the result. These tests require a doctor's diagnosis to determine to extent of the virus that has spread. There is a requirement of early detection of the covid-19 to save as many lives as possible. This proposed work proposes a comparative study between the different algorithms to apply the algorithm that fits the best. The comparative study helps to prove which algorithm gives the best accuracy for the particular problem. The proposed work includes aCOVID-19 prediction system which is automatic using the best fitting algorithm is used in places where experienced doctors and practitioners are unavailable, in this case workers in the field of health care can use this proposed work to prevent further loss of life by early prediction. Hence, a working methodology is proposed that predicts Covid-19 at early stages using a best fitting algorithm for the problem by training it with images of CT scan ofnon-Covid patients and Covid patients.

Keywords- Deep learning, ComputedTomography scan, LSTM, CNN, RNN, ANN, AUC, Segmentation.

INTRODUCTION

The covid-19 which is one of the most familiar and threatingword is a virus that spread a disease andmore contagious due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The observations made of having a COVID-19produce symptoms thatincludes cough, cold, high body

temperature, respiration issues, loss of taste and smell senses, headache, diarrhoea,sore throat, discolouration of fingers or toes. Covid-19 prevention measures are used to protecta person by avoid touching mouth,washing hands off, non-contact to the nose, face and eyes, maintaining a proper social distancingwith people and always wearing a face mask, when sneezing cover mouth and nose with a tissue or a bent elbow and getting vaccinated [1].

The transmission of this virus is through air borne means and droplets when exhaled by an infected person. When these particles are inhaled by a person through nose, mouth, eyes by touching, then the person becomes infected. The droplets of the virus can last from few minutes to few hours in the air [1]. The virus also lasts on difference surfaces for few hours to days depending on the surface that has been contaminated.The temperature and climatic conditions also impact the virus. Thevirus does not survive for a longer time in high temperature and humidity levels [2].

An early prediction of this virus gives a better chance of the patient to recover as early medical care is an efficient way to save the person. The earlier mentioned symptoms of covid are often variable which may begin one to fourteen days after the person is infected with the virus. In some cases, one third of the victims do not even develop noticeable symptoms after they have been infected. Whereas, out of these people who develop symptoms that are noticeable, 81% of these patients develop symptoms that are mild to moderate,14% of the patients suffer with severe symptoms and 5% of the patients suffer with fatality. The swab test takes time to receive the results, RT-PCR requires a day for the reports to be produce. This not only consumes more time but there is also a requirement of a doctor to help in the diagnosis. Hence, we propose a comparative study that helps choosing the most efficient algorithm. It shows which of the algorithm gives the best efficiency for an automated covid-19 prediction system [3]-[4].

The general detection of the virus is done through detecting the viral RNA through reverse transcriptase PCR (RT-PCR) of samples of the respiratory tract. However, this method has a lot of drawbacks: the turn-around time (TAT) ranges from a few hours to a few days which is relatively slow, this method has a low sensitivity which ranges from 59%-79% and it is expensive [5].X-rays and CT scans have been proved for more sensitive than RT-PCR in the detection of Covid-19. The image modalities that include Computed Tomography (CT)scan and commonly used X-Ray method taken over chest are non-interfering and used massively in health care for the detection of the severity and presenceof COVID-19.

A comparison between the use of CT scans and X-rays used in the detection of covid-19 shows its diagnostic accuracy. X-raysare not statically significant whereas, CT scans have better diagnostic performance over x-rays. Once the person has been affected by the virus, chest X-rays

remain normal even 4-5 days after the start of the symptoms but these symptoms cause noticeable changes in the CT scans. It has been reported that the diagnosis for X-ray is normal in both mild conditions and early stages. Visualization of internal structure in a 3D way can be seen in CT images can be considered as an effective tool in investigation. We use CT scans used for early prediction of covid-19 so the spread of the virus is curtailed [6].

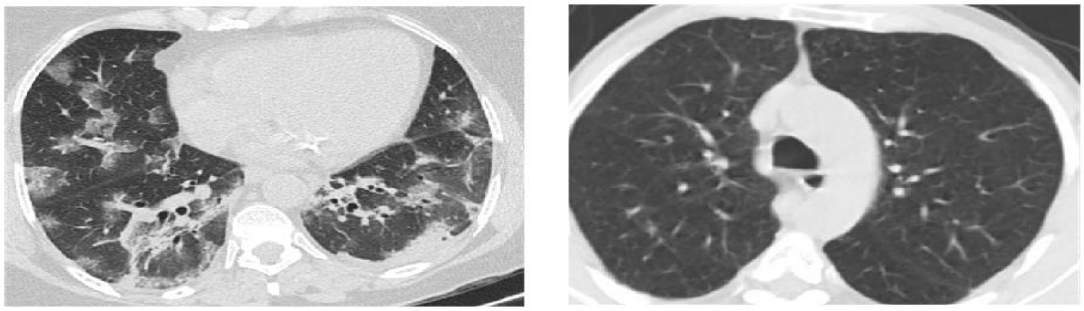


Figure 1: Example of (a) Covid-19 (b) non-Covid CT scan images in the left and right columns respectively.

Usability is often referred to as the measurement of the extent to which a specific user can use a specific product or design in order to achieve expected goal. Every system requires the function of usability. In this context, usability is combined with the medical field due to which extra care has to be taken as it involves the safety of a person. Radiologists' diagnosis has been critically used for detection of tumours and various other diseases. CT Scan usability has various barriers in relation to its diagnosis in the technical aspect. Hence, [7] focuses on resolving the technical barriers for the usability of CT scan. There is one radiology doctors and 11 CT Scan technicians interviewed in [7]. It highlights the importance of system engineers in the context of use. The system engineers have to develop a system in which the operational context to use has to be collected, analysed and stored. This is highly essential for the proper functioning of the system. Hence, we prefer using CT scans over the use of chest X-rays for the most suited algorithm in this study.

Motivation and Contribution of Work

The increasing cases of covid-19 and the fatality caused by them can be controlled and reduced by the early detection of the virus. The pandemic has affected the entire food system, caused restrictions on trade, affected farmers, jobs and livelihood of people. The restrictions on trading have prevented farmers and other industrialists to access markets which has in turn affected the lives of common people. The extreme rise in cases during the year 2020 caused lack of availability of hospital beds, oxygen cylinders, lack of medical staff that are specialized to handle covid patients,

shortage of other materials like PPE, ventilators, medicines and other medical supplies. The detection of the virus in the early stages is highly necessary to prevent the spreading to a larger mass of people. The contribution done to resolve the above stated problem is as follows:

- A throughout comparative study and analysis has been performed among all the traditional algorithms that are involved in detection done through CT scan images. This study enhances to implement a well-fitting algorithm that provides more efficiency and accuracy.
- A methodology for the prediction of COVID-19 cases from chest CT scan images of patientsat an early stage is proposed. We demonstrate the use ofCNN for solving the for novel COVID-19 prediction. We use the best algorithm on CT scans for predicting covid-19 cases.
- Two different kinds of CT scan datasets are deployed in the proposed produce the experimental results. The dataset of CT scan images has been split randomly for testing and training. Three different models of CNN have been implemented to show the accuracy these models provide to solve the particular problem that is put forward in this paper.

The organisation of this research work is such that the first section discusses the Covid-19 pandemic effect, causes and the repercussions caused by the pandemic, it emphasises on the background of prediction systems with their feature extraction process along with the motivation and contribution to carry out this work. The second phase consists of the already existing methodologies along with their short comings and various techniques that have been applied. The third section focuses on development of the mathematical model used for feature extraction and implementation process. The fourth section contains the results of the model that is implemented. This ends with a conclusion stating the outcome of this research.

LITERATURE REVIEW

A. LONG SHORT-TERM MEMORY (LSTM)

The analysis and work with deep learning applied to CT scans [8] has proven to be effective and efficient. There are huge amounts of data that has to be pre-processed and processed for proper and accurate results related to covid-19. Therefore, semi-supervised approaches of deep learning methodologies are in high demand which is proposed in [8]. The methodology that is proposed in this paper has an approach that incorporates the lung segmentation mask and also spatial and channel attention for feature extraction. These slices acquire axial dependency using LSTM. LSTM is defined as an artificial recurrent neural network (RNN) which in Deep Learning. There are feedback connections in LSTM andLSTM processes the all the data sequences and not only process single data points.

As a step of pre-processing, the methodologies of image enhancement on the basis of tone-mapping and stochastichave been evaluated for improving performance of the particular model. Lastly, the evaluation of performance done for the proposed framework has been conducted by the use of several configuration modules. The results that are obtained using LSTM applied in [8] to publicly available datasets show an accuracy of 81.9% and the average of Recall and Precision termed as the F1 score is 81.4%.

B. CNN AND LSTM

Cancer is the disease that spreads when the cells divide uncontrollably and it spreads to the surrounding tissues. Cancer has been the leading cause of dead and out of which lung cancer has been the second main cause of dead. The lung cancer detection at initial stages can help minimize the number of deaths caused by it. In the work [9], images of Computedtomography (CT) are deployed along with a system of Computer-Aided Diagnosis (CAD) systems. One among the effective tool used on medical imaging for depth medical analysis is CAD. The main focus in this paper is the development of a CAD system that is advanced by the use of deep learning algorithms. This developed system extracts the necessary information from the provided images of CT scan and performs lung cancer diagnosis in a timely manner. [9] has three parts that is namely, segmentation, feature extraction and classification.OSTU Thresholding is used in [9] to perform automatic image thresholding used to segment the images of CT scan. CNN model is used in feature extraction and RNN- LSTM is used in classification and obtaining high accuracy. Therefore, [9] proposes a hybrid system of CNN-LSTM for lung cancer detection. The CNN model performs feature extraction whereas LSTM performs the classification. In this paper, comparison is performed which compares the CNN-LSTM model to the existing models based on the accuracy measure. It mainly aims on improving the accuracy of prediction systems.The system provides a 97% accuracy which is the highest achieved for the diagnosis of lung cancer.

[10] mainly focuses on automated detection of brain haemorrhage from CT scan images. A brain haemorrhage is normally caused due to high blood pressure and trauma. An artery in the brain bursts and this causes bleeding within the brain. This normally has a fatal impact on the functioning of the brain. The normal detection of brain haemorrhage using CT scan images is normally done by professional radiologists. The model in [10] stimulates the analysis of radiologists in CT scans that is in 3D in the real world. In order to incorporate 3-D inter slice context, aslice level classification combination of task as well as sequence level task is used to model a 3D labelling task. In [10], usage of ConvolutionalNeural Network (CNN) combined over network of Long Short-Term Memory (LSTM). Also, an architecture ofDenseNet is used for CNN to learn slice level features. Neural network uses a DenseNet architecture for recognition in visual object. There is a similarity

ofDenseNetandResNetthat has some differences that are fundamental. For the combination of spatial dependencies between slices we use bidirectional LSTM. Prior information about haemorrhage region is also used to focus DenseNet towards the relevant features. The modified architectural model that is used in this paper is termed as Recurrent Attention DenseNet (RADnet).

C. CNN

The coronavirus has caused a huge impact globally and millions of lives have been lost. The countries with weaker healthcare systems have suffered enormous losses. Deep Learning in general has been implemented in healthcare such as in diabetes detection, cancer detection, lung module classification, etc. [11] a survey is done which incorporation of deep learning methodologies in the field of healthcare for better detection and treatment of covid-19 cases. The normal detection of disease is generally done through medical imaging (e.g., X-rays, CT scan, MRI), when this is combined with deep learning techniques then there is a more efficient and faster way to treat the patients that are suffering. [11] mainly focuses on the state-of-the-art works that are relating to deep learning methodologies. The study shows the most efficient of the algorithms to be Convolutional neural networks.[11] also list few limitations: the outbreak of data relating to covid datasets has led to a complex pattern and also shows an extreme variation in the behaviour, obtaining largenumber of covid-19 datasets that are high quality is very difficult considering the privacy of the patient, each country has different regulations containing the sharing of covid-19 data.

Automated detection of lesions, organ detection, radiotherapy, uses organ segmentation or annotation. [12] focuses on using Convolution Neural Network (CNN) for the segmentation. This CNN is a model of machine learning that requires large amount of training data for particular image analysis. Medical image analysis is normally time consuming and expensive hence, here we use crowdsourced kidney segmentations. It evaluates the efficiency and accuracy of crowdsourced kidney segmentations. Crowdsourcing has been considered an effective method to develop large amount of data for machine learning purposes and they are also used in the field of medicine. The possibility to crowdsource large amounts of CT scan annotations with segmentations for every organ, this technique drastically changes the training data that is available for CAD.

1. VGG16

Convolution Neural Network (CNN) is a class of Deep Learning Neural Network. CNNs are various versions of multilayer perceptron. Multilayer perceptron is a feedforward ANN. The use of CNN models in detection using image analysis is more efficient. The conventional testing methods have proved to be slow and the number of testing kits are insufficient. It makes sense and essential to have accurate and fast detection of the covid-19 virus.The use of Computed Tomography (CT scans)

for the covid-19 detection is shown in [13]. The use of CNN models requires a large amount of data. The use of massive and huge datasets using ImageNet.In ImageNet [13], there are models that are pre-trained which adapt to the tasks of medical imaging. This makes the performance better, increases computational speed which makes it time efficient. There are three CNN models that are used in [13] for the initial covid-19 virus prediction. These models implemented in [13] include Xception, VGG19 and VGG16 are pre-trained using ImageNet. The highest accuracy, sensitivity and F1 score was of VGG16.

There is a requirement for doctor to make decisions medically on Covid-19 based on examinations on their patients. However, in [14] transfer learning is used in several researches. One such machine learning methodology used is Transfer Learning in which a model for has been reused for a task is developed at starting point for a second task of the model. It focuses on a single modality of biomarkers fordiagnosing Covid-19 pneumonia. The detection of this virus by X-ray and CT-scan of the chest imaging. In [14], two different transfer modelsis combined and used. 2500 X-ray images as well as 2500 CT scan images are used for classification of two classes: normal conditions as well as Covid-19 Pneumonia. In this paper, a concatenation of models VGG16 and ResNet50 is used, this shows the maximum accuracy of 99.87%. There is also a comparison of CT-scan ResNet50 networks in single modality which gives a 98.00% accuracy.

2. DENSENET AND U-NET

An important biomarker of cardiovascular diseases is the quantityof deposits of calciumin coronary arteries. A CT scan of coronary shows calcium deposits to detect extreme cardiac problems also termed as artery diseases. [15] proposes automated system to calculate the coronary calcium score using chest CT scans is developed. The main complexities caused in our algorithm were(a) confusion between calcifications in the aorta (b)false-positive calcified mitral valve and the arteries that are close to aortic wall. Two network architectures have been evaluated. These two network architectures provide the highest accuracy.

3. ALEXNET

The most critical step in the coronavirus research is the diagnosis of the virus. This virus has similar exhibitions to various pneumonia types. The system proposed in [7], there are two algorithms of optimizations used in classification and feature selection of COVID-19. The features that are extracted from [16] using a Convolutional Neural Network that is AlexNet. The proposed model uses two datasets to be tested: CT imagesthat result as Covid positive and CT images that result as Covid negative. The proposed algorithms has been improvised as compared to the various algorithms that involves the algorithms of original WOA, to calculate the minimum value of fitness for feature

selection of the features that have been extracted from the datasets of Covid-19.The algorithm that has been proposed for feature selection (SFS-Guided WOA) has been compared amongst various other algorithms that has been used for the literature analysis for efficiency validation. In [16], there exists a voting classifier that is proposed (PSO-Guided-WOA) which is evaluated using 0.995 AUC making it much efficient in comparison to the other voting classifiers while consider the metrics for performance.

The Centreline Spine System is a procedure that reduces any disruption to the nerves, tissues and muscles that surround the surgery site.[17] uses the CT scan of the spinal cord and proposes a system that is fullyautomated with the concatenation of two CNN networks along with an algorithm for spine tracing. A special CNN, called SpineCNN is used which performs vertebrae segmentation. The recognition and detectionof imaged spine segment at the ends are performed with the Spine CNN in [17], this is fine-tuned using AlexNet. A population approach is used which yields more distinct and efficient results as compared to using any other algorithm.AlexNet is chosen as the most efficient model because although it provides closer accuracy as the other models, it has the lowest computational time in comparison. The method used in [17] gives an accuracy of 89.4%. This also shows that simplest of algorithms can give the best results for a common classification problem.

4. MODIFIED ALEXNET (MAN)

Lung abnormalities are normally detected through CT scans and x-rays by expert radiologists. The commonly found abnormalities in the lung are bronchogenic cysts, congenital pulmonary airway malformations, congenital lobar overinflation, congenital high airway obstruction syndrome, bronchopulmonary sequestration. If these abnormalities are detected at an early stage, then the early treatment of the patient can begin thus saving the lives of people. [18] proposes a framework in Deep Learning to focus on cancer and lung pneumonia. It uses a Modified AlexNet (MAN) framework in deep learning to detect the abnormalities in the lung. A filter is used to omit the artifacts in the CT scan images of lungs. In [18], EFT is used which is Ensemble-Feature-Technique, this is done by integration of the handcrafted features and deep features. A classification accuracyof around 86.47% is achieved by the MAN that is proposed along with the support vector machine (SMV)classifier. And with the EFT the deep learning framework provides an accuracy of >97.27%.

5. GOOGLE NET

The early detection of covid cases has become an immediate need considering its impact on the health and economy globally. The most efficient method for this early detection would be through the use of X-rays and CT scans. Although the use of CT scans is more sensitive and detect symptoms

at the earliest stage, X-rays are easily available for use and are used in the approach in [19]. The method in [19] uses GoogleNet to predict Covid-19 by the use of chest X-ray images of patients. GoogleNet is a type of CNN architecture which is deep by 22 layers. [19] uses a transfer learning model with GoogleNet. The GoogleNet network learns a feature representation that is rich for a large and widerange of images. This model when trained obtained a 99% training accuracy and the testing accuracy was said to be 98.5%.

D. R-CNN AND KNN

Stroke occurs when blood supply is cut off from a part of the brain and the brain stops working. In turn, the part of the body that is controlled by that part of the brain is damaged and stops working. This affects almost 16 million people around the globe, among this 38% of the cases land up in fatalities. Medical diagnosis of stroke is done super effectively by computed tomography (CT). There are a lot of variations in the analysis for the perception charcteristics in specialists. The system in [20] proposes an automated system Health of Things that are capable of classification of CT images through networks of deep learning in the skull. Machine Learning methods are combined with CNN thatsegments stroke along a transfer learning process. Accuracy attained in classifying images, between (haemorrhagicstroke or non-Injured), andspecificityof 99.93% along the best model of segmentation (Mask R-CNN + KNN). Segmentation had a 4 second response time, this proves the method in [20] is efficient and robust.

E. RESNET

A simple conceptual framework is made specifically for rapid COVID-19 screeningwith 3D CT images of chest is proposed in [21]. The framework that is proposed in [21] can efficiently predict if a CT scan of a patient is positive for pneumonia and it also simultaneously identifies the various pneumonia types that are present in ILD (Interstitial Lung Disease) and Covid-19. It has anintegration of two branches on the basis ofResNet into a single framework model fortraining by design a prior-attention residual learning (PARL) block. ResNet also called Residual Network was introduced after Convolutional Neural Network. It is normally used for many tasks of computer vision. Within these blocks, the information on hierarchy from the detection of lesion region branch which has been transferred to the classification branch of COVID-19 fordiscriminativerepresentation that has to be learnt.

Early detection of covid-19 pandemic disease causes the spread to be curtailed. The new variant in Covid-19 virus which is the pneumonia type coronavirus has caused uncountable number of deaths across the world. The normal testing of coronavirus takes time, whereas diagnosis through CT scan images can be done faster. But the diagnosis through CT scan images can be done by expert

radiologists who have to look at the numerous slides that areseen in the images of CT scan. The lack of radiologists can also cause a delay in the treatment for patients that are affected with covid-19. The use of artificial intelligence inthese cases could cause a more rapid and efficient diagnosis. In [22], artificial intelligence used in the diagnose through CT scan images for classifying patients having covid-19 and normal CT scan images. The method that is used in [22] is ResNet-50 model of deep learning for the prediction of covid-19 with the use of 3D CT scan images. Image-level prediction is used in AI that provides a 96% AUC of covid-19 detection cases amongst normal images of CT scan.

The automated covid-19 prediction cases are a better and faster way for diagnosis which is done using CT scan images. In [23], a new dataset that has a total of 48,260 images, in which the images of normal people are 282and 15,589 images taken from 95affected patients that have covid-19. [23] proposes an algorithm of image processing, whichfilters the patient'simages that properly show the lung, hence increasing the speed and the network accuracy. A Convolutional Neural Network (CNN) for the better improvement and to proceed further at the classification phase was introduced. The model uses ResNet50V2, this architecture achieves the best results. After the model has been trained using these training datasets, the trained network is used to run the completely automated covid-19 identifier system. The model showed a 98.49% accuracy for a single classification image system.

F. ANN

Cancer is a disease that involves abnormal growth of the cells, these cells divideuncontrollably and this spreads to other parts. There are four types of cancer Melanoma,Carcinoma, Sarcoma, Lymphoma andLeukaemia. Around 18 million new cases of cancer occur annually. Many of these cancer cases usually end up fatal. One such way to prevent the fatality of this disease is the earlyprediction and treatment of cancer. In this way there is a higher chance that the person survives. [24] proposes a method for identifying cancer from the lung images of Computed Tomography (CT) using Artificial Neural Networks (ANN). There is segmentation of the whole lung from the images of CT scan and parameter like mean fifth central movement, standard deviation, sixth central movement, kurtosis and skewness is calculated and they are used for classification. The feed-forward back and feed-forward propagationof neural networks have been deployed for classification. Comparatively, feed-forward back propagation method is more efficient and amongst its thirteen of its training functions, Traingdx function produces an accuracy of 91.1% which is the maximum. In [24] thereare two new training methods that are proposed. Where the first training function produces an accuracy of 93.3% as well as an error of 0.998, while the second training function produces a 91.1% accuracy and an error of 0.0942.

G. FRACNET

Diagnosis of fractures in the ribs has an important function in the identification of trauma severity. However, the quick and precise detection and identification of rib fractures having major numbers of images in CT scanwhere the count of patients is also constantly increasing requires a qualified radiologist. A deep learning system is proposed in [25] for the detection and segmentation of the rib fractures from images of CT scans.Thin-slice using images of CT scans of a high-quality with a thickness of 11.25 mm have been used in this study. It has been reported through analysis that images of thin-slice could provide diagnosis of incidental findings as well as bone fractures. [25] shows the development of a model of deep learning, called asFracNet, for detection and segmentation of rib fractures.Evaluating the sensitiveness and the false positives for the performance detection is done by the use of FreeResponse ROC (FROC) analysis. The Dice Coefficient (Dice) and Intersection-over-Union (IoU) is involved in the evaluation of segmentation performance of the predicted rib fractures. This method used in [25] helped achieve a detection of 92.9% for sensitivity along with 5.27 false positives for every scan and a 71.5%segmentation Dice on the test cohort.

H. COVTANET

The fast and rapid spread of the virus is a major challenge that is faced to control the spreading virus. Aneural network that is hybrid has been proposed in [26] which is named as CovTANet, atool is provided by this network which helps in the early diagnosis, prediction of covid by the use of chest CT scans and lesion segmentation. A strategy based on multi-phased optimization is used in [26] to solve the challenges of diagnosis that are complicated at the initial stages of the infection. Initially, a network for lesion segmentation is optimized. Furthermore, integrated in a joint optimization to predict and diagnose whichprovides feature enhancement for regionsthat are infected. The CovTANet that is proposed in [26] gives an accuracy of 85.2% for the covid patients that have even mildsymptoms and it shows an accuracy of 95.8% for patients with severe cases.

I. WATERSHED ALGORITHM

The food that is consumed in today's times causes a lot of problems in all age groups. Mainly, fatty liver and the consumption of alcohol also causes cirrhosis, many other liver diseases are also caused such as hepatitis A, hepatitis B, hepatitis C, jaundice, liver cancer and even liver failure. The symptoms of these include loss of appetite, pale stool colour, swelling in legs and ankles, abdominal pain and swelling, skin and eyes appear yellowish, chronic fatigue, tendency to bruise easily, itchy skin and dark coloured urine. All of this reduces the functioning of the liver, affects thehormones, nutrients and protein production in the body [27]. The normal way of finding these abnormalities in

the liver is by doing a CT scan of the abdomen. Watershed transform algorithm proposed in [27] is mainly a segmentation algorithm and is used in [27] for the segmentation of the liver; this is used to produce locations of the liver that distinguishes objects for the background. The scan is further segmented by a method of binary threshold which separates image of liver as the object observed. The area that is affected is finally calculated. [27] states that the wide segmentation that is done with a81.15% accuracy and the segmentation with disease having 98.28% average accuracy.

J. SEGMENTATION ALGORITHMS

A brain haemorrhage, also known as cerebral haemorrhage is a kind of stroke. This is normally caused due to high blood pressure and trauma. An artery in the brain bursts and this causes bleeding within the brain. This normallyhas afatal impact on the functioning of the brain. At times, the identifying brain haemorrhage immediately is not possible. The normal identification of brain haemorrhage is done with professional analysis of CT scan. To ease this diagnosis there is a requirement of segmentation of CT scan which has to be automated and quickly done. If there is accurate and fast segmentation of brain bleeding the patients suffering can be diagnosed fast and obtain medical treatment immediately [28]. These CT Scan images in [28] are first pre-processed in the first phase which is the pre-processing phase which uses colour filtering, erosion and dilation methods. This phase is mainly performed to check for cerebral haemorrhage and to eliminate any noise that is present in the CT Scan image. After this phase, a cropping segmentation is performed to separate the skull bones and the brain tissues. Median filtering is used to improve the quality of the image, after which the image is further segmented in order to separate the image from cerebral haemorrhage. Then the percentage bleeding of the brain is calculated. There is an average error of 11.17% in the test calculation of the area of bleeding of the brain.

K. CLAHE

When the supply of blood is blocked or reduced to any part of the brain, this prevents the brain from getting oxygen and the necessary nutrients. Due to this the brain cellsrapidly begin to die. The symptoms of stroke include paralysis of the face, the arm and leg,trouble to walk, understand and speak. Basically, this is a harmful disease that brings upon physical impairment on the patient. The diagnosis of this is normally done using CT scans but it is not always efficient since there are lot of parts in the CT scan that are unclearand invisible. This requires the image quality of the image to be improved. An adaptive histogram equalization (AHE) technique that can cope to enhance the contrast in local areas. Since this enhancement can sometimes be excessive hence, we use (CLAHE) which is Contrast Limited AHE. Therefore, [29] focuses on enhancement of the images using CLAHE for strokeidentification in order to distinguish the CT scan as a normal one or with stroke.

PROPOSED WORK

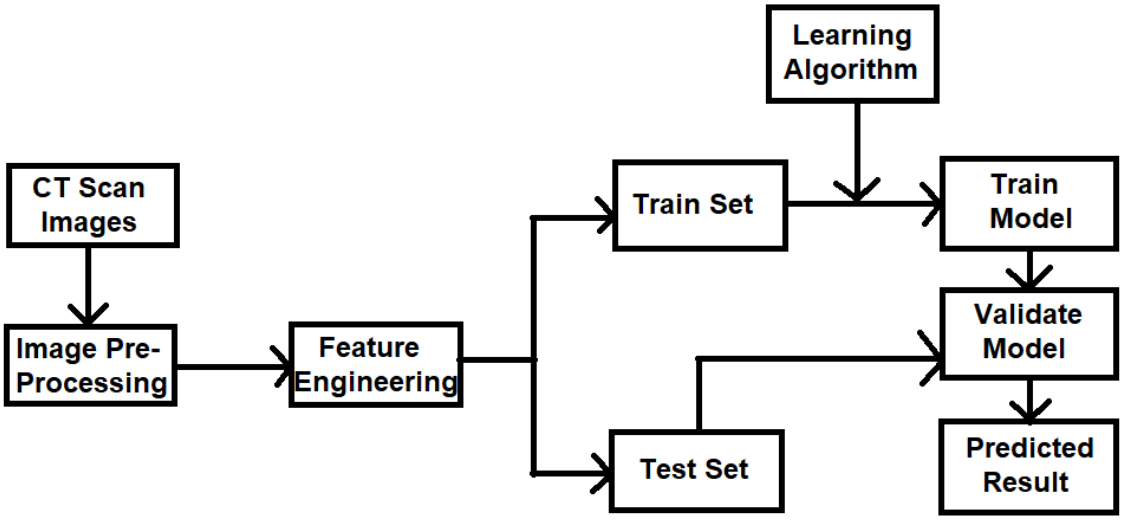


Figure 2: System Architecture of the Proposed System

The figure 2 shows the overall system architecture of the proposed system. Firstly, the images of CT Scan are collected of both non-covid and covid patients. Once the images are collected, they are pre-processed further for feature selection. The image that is under the data set of covid and non-covid are studied and the value of every pixel is plotted on a histogram for better clarification. Each of the values in the image are selected to be studied by the system. The dataset is then split randomly into test data and train data. Into which the algorithm is run with the train dataset. In this proposed work, there are three models, namely: Alexnet, Modified Alexnet and Resnet which are used. The train dataset is used in training each of these models individually. The test dataset is than used for validation. In the validation process, both the test and train dataset are used. The accuracy, precision, recall and F1 score is determined for each of the models. A confusion matrix is also been plotted to show the values that have been developed by the models. Each model shows its own accuracy using the same train and test dataset.

The model that is used in the implementation of the early prediction system is Convolutional Neural Network (CNN). There are namely three types of models that are selected for proposed methodology by careful analysis performed by a comparative study that is done above. An analysis on the basis of comparison is performed among the three models is performed using same datasets and these metrics evaluations are considered to find the more efficiently working model. The experimental results are demonstrated which indicate performance efficiency.

Convolutional Neural Network (CNN)

Convolutional Neural Network is a deep neural network. CNN is a feed forward network which can extract properties from an image and process pixel data. CNN model is mainly used for image processing, it has been designed to recognise visual patterns from the pixel images with the least amount of pre-processing. Out of the entire survey that is done considering all the algorithms that are used on various healthcare system with CT Scan images as data set [30].

Alex Net

Alexnet is a type of Convolutional Neural Network (CNN) that is used in this proposed work. The network consists of several convolution layers that are combined with pooling layer and nonlinear layers. The CT scan images that are used in one layer as input pass through the layer and the output of that layer becomes the input for the next layer. This process continuous for every layer of the model. A nonlinear layer is added after each convolution operation. This layer consists of an activation function. This layer is what makes the convolutional network more intense and makes it operable as a class label. This nonlinear layer is then followed by a pooling layer. The pooling layer along with the width and the height of the image and performs operations on the images. This causes the reduction of the image volume. If some features that are unwanted are identified then the dropout operation is performed at this layer. This causes the load on the model to decrease since the unwanted features are dropped and not considered. These features that are considered in the previous operation are no longer considered in the next operation. After all these layers undergo the processing a fully connected layer is added.

Res Net

ResNet stands for Residual Neural Network which is type of artificial neural network that is used in deep learning and machine learning applications. This model is basically used for computer vision tasks. ResNet was introduced after the Convolutional Neural Network, the accuracy and performance has been increased and elevated by the introduction of ResNet. Every layer that is added in ResNet is used for identity mapping, this model is implemented with double or triple layers by skipping the nonlinear layers in between them. Skipping layers in the middle simplifies the model and makes the understandability of the model easier. This presence of fewer layers in the model increases its speed since there are fewer layer that have to be processed to produce an output. The network then gradually replaces the layers that have been skipped by learning the feature space.

Modified Alex Net

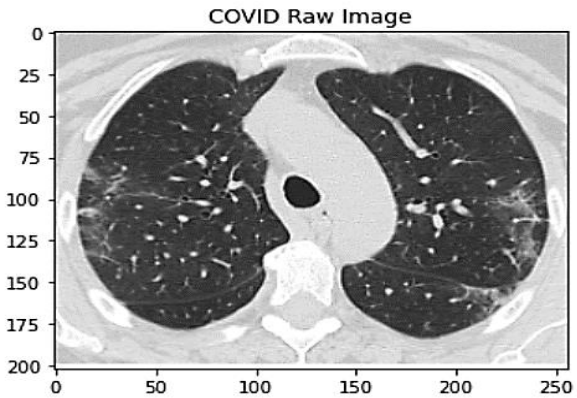
This model is a Modified Version of AlexNet which is normally used in many image processing applications. This model is used for adaptiveness towards various sizes that are caused due to image resolution as well as perspective. Feature maps of various CNN layers are used and this model is used to cope with various image block sizes.

Details of Dataset

The dataset used for this implementation is run through three different types of models that are selected by detailed analysis of the study that is done above. The data that is used in this proposed work consists of two different categories. These categories include Covid dataset and non-covid dataset. The models that are used in this paper use both the datasets. The dataset is split into two parts, one part for training and the other part for testing. The system uses a random splitting method that splits the data 80% for training and 20% for testing purposes.

Feature Extraction Process

The image dataset is pre-processed and then the feature extraction process takes place. It is completely different how we see the image and how to computer sees the image, the computer sees the image in an array of pixels. For example, if the original size of the image is 300 x 300, the computer perceives the image as 300 x 300 x 3. Where the first 300 depicts the width, the next 300 represents the height of the image and the last 3 represents RGB channel values. The computer system has been assigned various values from 0 to 225 for each of these numbers. These pixel values indicate the intensity of the pixel at each point. For the problem that has been proposed in this work, we solve it using the characteristics at the base level. In the understanding of human, these features are perceived as either size of the alveoli or fluid in the alveoli but the computer sees it in the form of curvatures, graphs and boundaries. Later using the various layers of CNN, we construct more abstract models using this as the base.



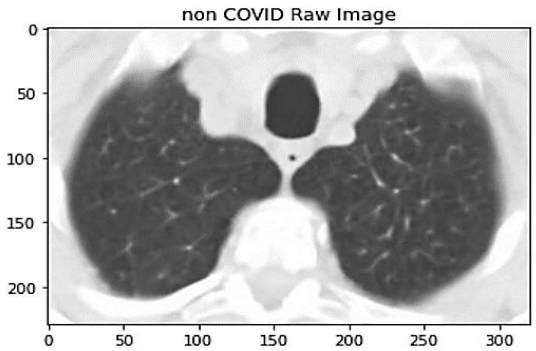


Figure 3: (a) Raw image from Covid Dataset (left) and (b) Raw image from non-Covid Dataset (right)

The above figure 3 shows the CT scan image of covid and non-covid data. These images are used to observe the features that are required for predcition of the disease in the lungs. For a better identification and distinction of the covid and non-covid images. The use of a histogram is to graphically conclude the univariate distribution of a dataset. A histogram is plotted for a better understanding of the images by a graphical curve which is shown in figure 4. It shows the distribution of skewness and the gaussian curve. Skewness is the measurement of the asymmetry in the histogram which is the frequency distribution. A guassian curve is the probability distribution which appears to be symmetric about the mean. The system percieves these details in a more improvised manner when graphical methods are used.

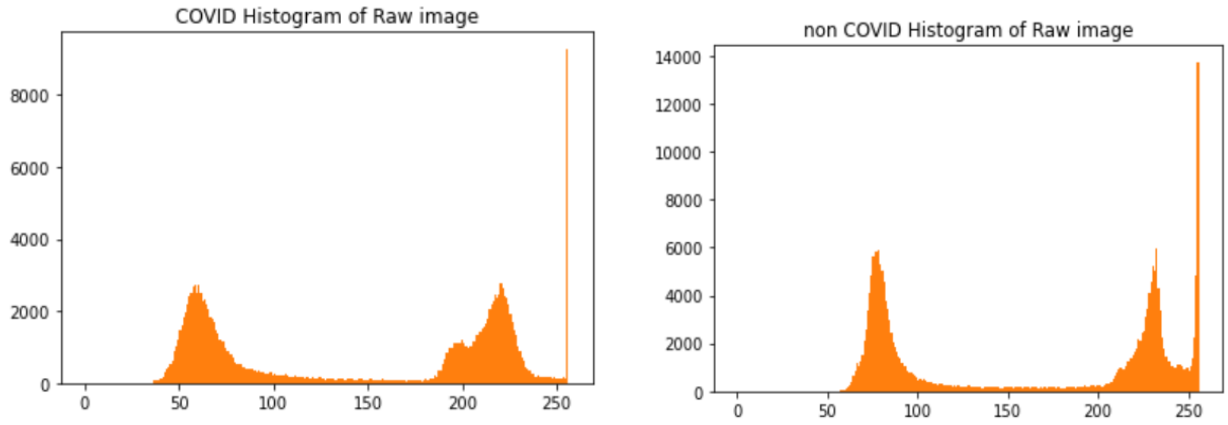


Figure 4: Histogram of (a) Covid raw image (b) non-Covid raw image

The figure 5 shows histogram equalized images of both covid and non-covid datasets. The histograms that are equalized also called as histogram equalization a technique in computer image processing which is used in improving the image contrast to make images more distinguishable. This is done by spreading out the most frequent intensity values. The intensity range of the image is stretched. The histogram equalization helps the areas of the image that have a lower contrast to gain a higher contrast. This is shown in 6, where the value of frequent intensity at maximum are spread out for better and more clarified distinction.

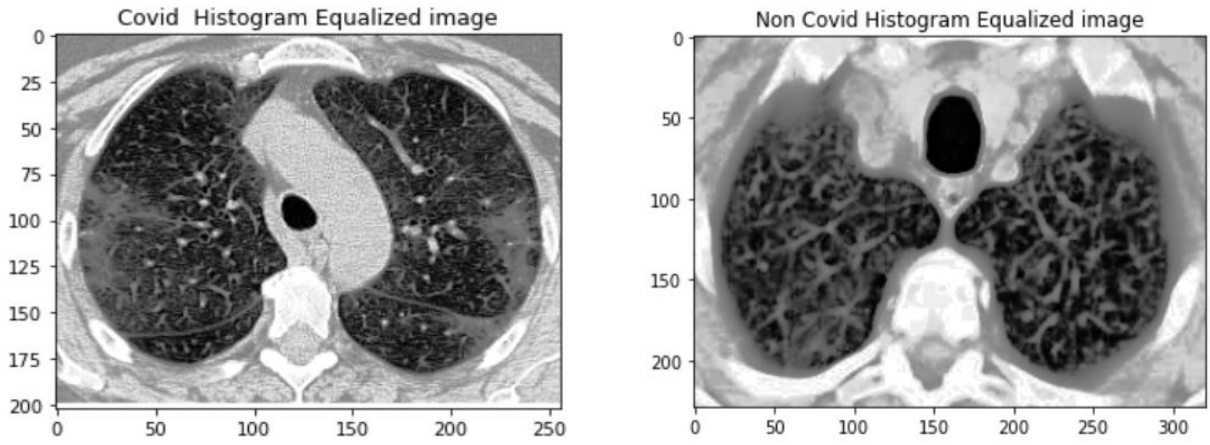


Figure 5: Histogram Equalized Image of (a) Covid Histogram (b) non-Covid Histogram

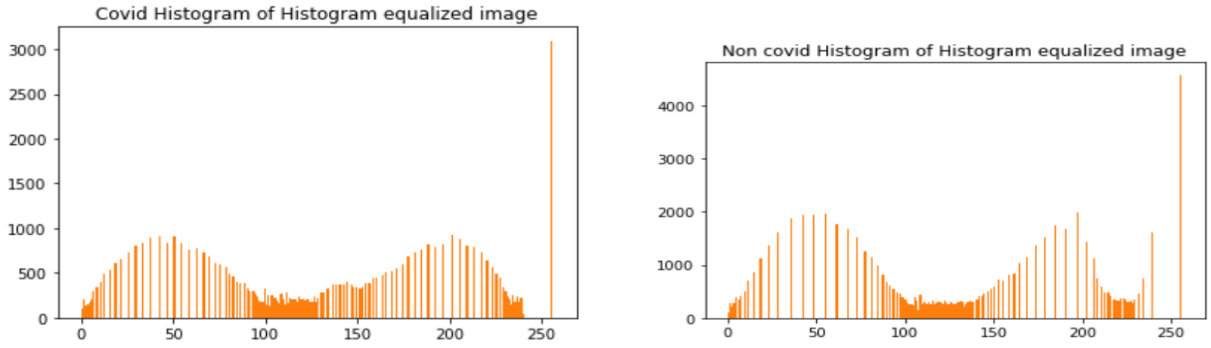


Figure 6: Histogram of (a) Covid Histogram (b) non-Covid Histogram

After the histogram equalization of all the images in both the categories of datasets the limit for contrast is set and the three models of CNN are constructed for the prediction process. These images are then passed throught the models after which the performace matrices such as accuracy, precision, recall and F1 score are calculated.

Results

All the three models belong to the CNN algorithm as the highest efficient that is produced is for this algorithm on CT scan applications, the results produced are highly efficient and accurate. The model performance for early prediction is measured using various performance metrics such as precision, accuracy, recall and F1 score, these metrics help in performing a comparative analysis among the models that have been used in this research.

Accuracy of Algorithms

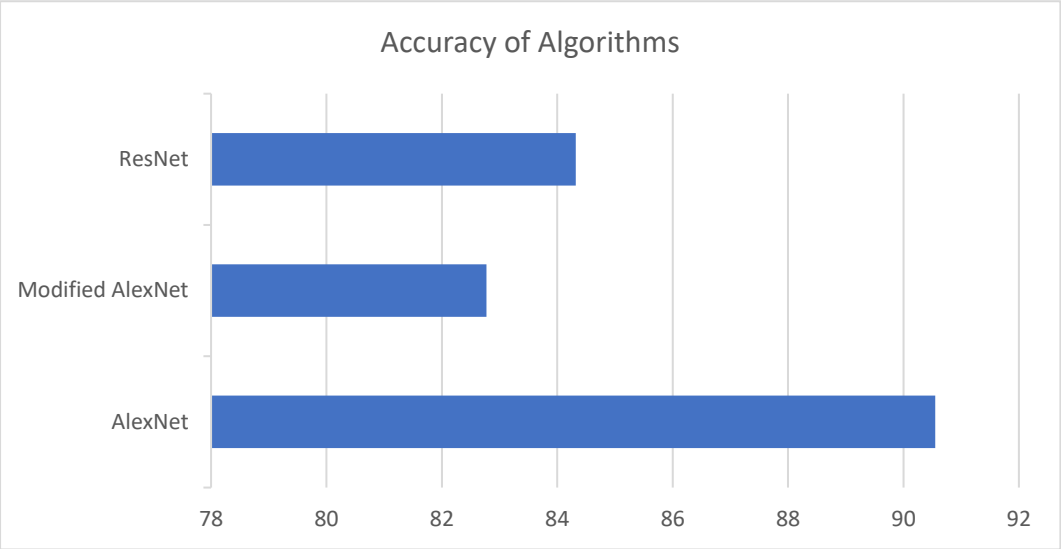


Figure 7: Accuracy of proposed algorithms

The figure 7 shows the accuracy of the algorithms. From the figure, it is proved that Alexnet has a maximum accuracy, followed by ResNet and lastly Modified AlexNet. These three algorithms have the most accuracy for covid-19 virus prediction at anearly stage. The table 1 below gives the exact accuracy values of the three models of CNN.

Table 1: Accuracy Values of the Algorithms Precision of Algorithms

Algorithm	Accuracy
AlexNet	90.55
Modified AlexNet	82.77
ResNet	84.32

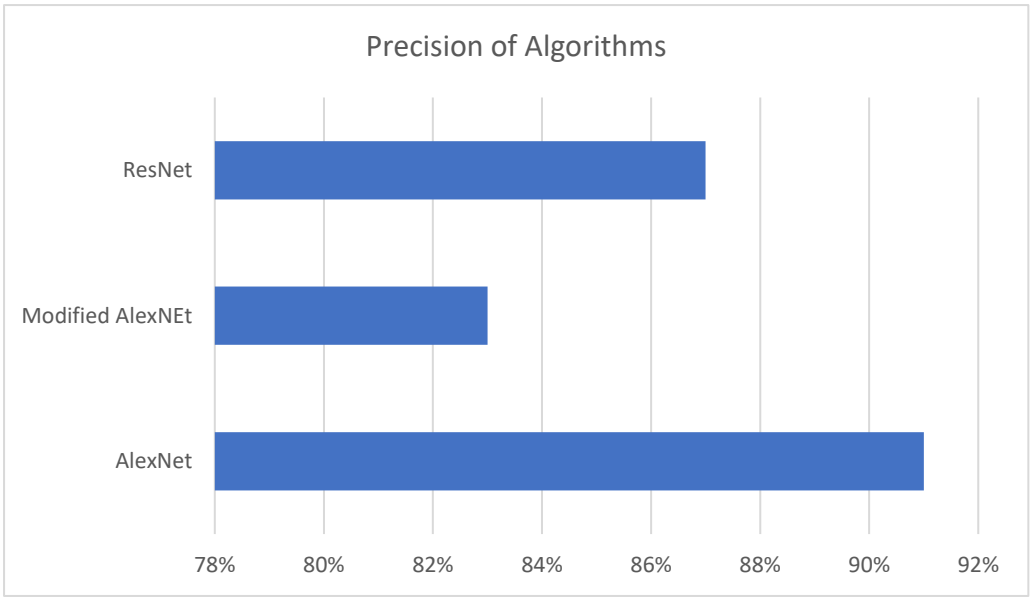


Figure 8: Precision of proposed algorithms

Precision in a model refers to the ability of that model of identifying only relevant and necessary points or features needed to achieve higher accuracy. Figure 8 shows the precision of the three models of CNN. Firstly, AlexNet shows the highest precision, followed by ResNet and lastly Modified AlexNet. The table 2 give the values of precision for the models that are proposed.

Table 2: Precision Values of the Algorithms Recall of Algorithms

Algorithm	Precision
AlexNet	91%
Modified AlexNet	83%
ResNet	87%

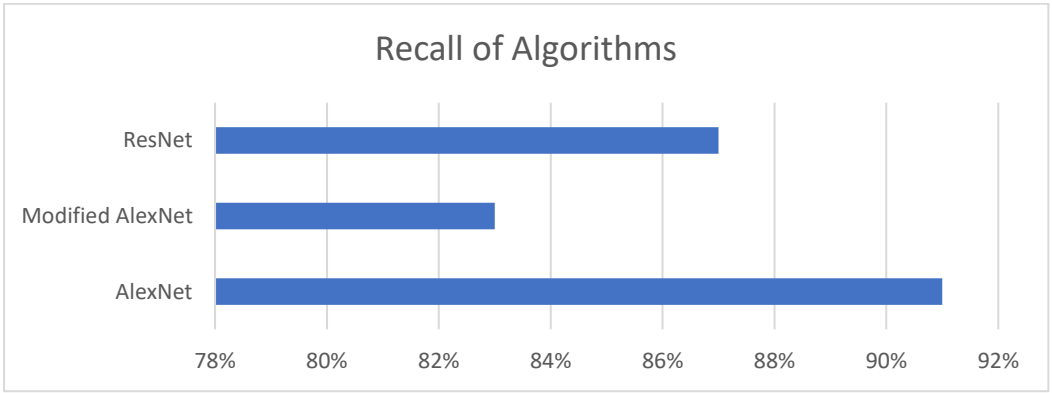


Figure 9: Recall of the proposed Algorithms

The measurement of the model to correctly identify the true positives is termed as recall. The figure 9 states the recall of all the models that are implemented in the proposed work of this paper. AlexNet has the highest recall, which is followed by ResNet and lastly, Modified AlexNet has the least recall value. The exact values of recall are stated below in table 3.

Table 3: Recall Value of the Algorithms Score of Algorithms

Algorithm	Recall
AlexNet	91%
Modified AlexNet	83%
ResNet	87%

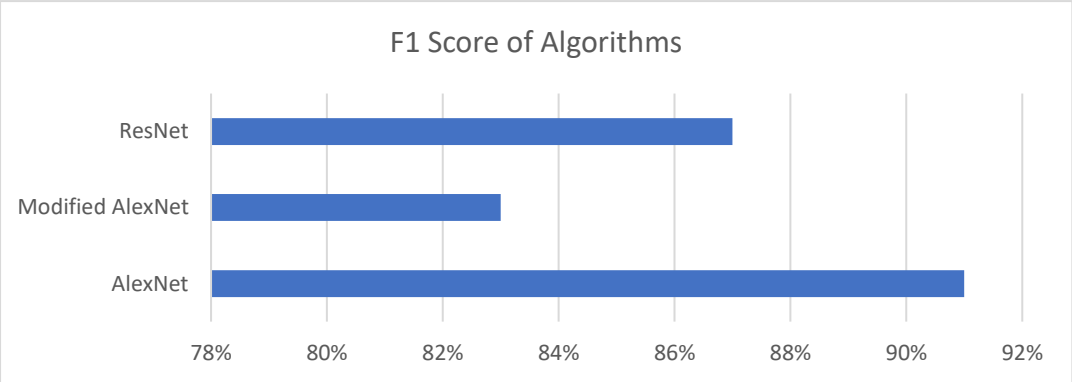


Figure 10: F1 Score of the proposed Algorithms

F1 score also known as F score or F measure is used for conveying the balanced between the recall and the precision. The F1 score of the algorithms proposed are given in the figure 10. AlexNet has the highest F1 score which is followed by ResNet and the least F1 score is of Modified AlexNet. The specific values of F1 score are listed in the table 4.

Table 4: F1 score values of the Algorithms

Algorithm	F1 Score
AlexNet	91%
Modified AlexNet	83%
ResNet	87%

Conclusion

The significance and importance of analysing and implementing early covid prediction is implicated in this research work. The feature extraction process occurs by which three models are used to

perform the prediction. These three CNN models are selected to be the best fitting models after a detailed comparative study and analysis is performed on all the algorithms that use CT scan images for prediction methodologies. It is proved through experimental results that out of the three models used in this paper: AlexNet, ResNet and Modified AlexNet, the AlexNet model is found to be the model with the highest accuracy and works most efficiently.

References

[1] Tawsifur Rahman, AmithKhandakar, Md Enamul Hoque, Nabil Ibtehaz, Saad Bin Kashem, ReehumMasud ,LutfunnaharShampa , Mohammad Mehedi Hasan , Mohammad T. Islam , Somaya Al-Madeed , Susu M. Zughaier , SaifBadran , Suhail A. R. Doi ,Muhammad E. H. Chowdhury, (2021), "Development and Validation of an Early Scoring System for Prediction of Disease Severity in COVID-19 using Complete Blood Count Parameters", IEEE Access, Volume XX, 2021, DOI 10.1109/ACCESS.2021.3105321.

[2] Ping He, "Study on Epidemic Prevention and Control Strategy of COVID-19 Based on Personnel Flow Prediction", 2020 International Conference on Urban Engineering and Management Science (ICUEMS), DOI 10.1109/ICUEMS50872.2020.00150.

[3] Haiying Ren; Jianfeng Shen; Xiaoyong Tang; Tianyi Feng, "5G Healthcare Applications In COVID-19 Prevention And Control", 2020 ITU Kaleidoscope: Industry-Driven Digital Transformation (ITU K), doi: 10.23919/ITUK50268.2020.9303191.

[4] Dicky Dwi Putra, Mohammad Febriyanto, Muhammad Miqdad Nadra, WervyanShalannanda, Elsa RamadhaniFirzal, Achmad Munir, "Design of Smart-Gate Based on Artificial Intelligence Possibly for COVID-19 Early Prevention at Public Area", 2020 14th International Conference on Telecommunication Systems, Services, and Applications (TSSA), doi:10.1109/TSSA51342.2020.9310878.

[5] Kang-Yi Lien, Wang-Ying Lin, Chih-Hao Wang, Huan-Yao Lei, &Gwo-Bin Lee. (2007). "Miniature RT-PCR systems integrated with a sample pretreatment device for virus detection" 2007 IEEE 20th International Conference on Micro Electro Mechanical Systems (MEMS). doi:10.1109/memsys.2007.4433018

[6] Prerak Mann; Sahaj Jain; Saurabh Mittal; Aruna Bhat, "Generation of COVID-19 Chest CT Scan Images using Generative Adversarial Networks", 2021 International Conference on Intelligent Technologies (CONIT), doi: 10.1109/CONIT51480.2021.9498272.

[7] Saad Aldoihi, Omar Hammami, 2020, “Factors Contributing to CT Scan Usability”, 2020 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), doi: 10.1109/CIVEMSA48639.2020.9132747

[8] Mohammed Ahmed, Wang Congcong, Zhao Meng, Ullah Mohib, Naseem Rabia, Wang Hao, Pedersen Marius, Cheikh FaouziAlaya (2020), “Semi-supervised Network for Detection of COVID-19 in Chest CT Scans”, IEEE Access, 1–1. doi:10.1109/ACCESS.2020.3018498

[9] Mhaske. D, Rajeswari. K &Tekade. R, (2019), “Deep Learning Algorithm for Classification and Prediction of Lung Cancer using CT Scan Images”, 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA). doi:10.1109/iccubea47591.2019.9128479

[10] Grewal M, Srivastava M. M, Kumar P &Varadarajan S, 2018, “RADnet: Radiologist Level Accuracy using Deep Learning for Hemorrhage Detection in CT scans”, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). doi:10.1109/isbi.2018.8363574

[11] Bhattacharya S, Reddy Maddikunta P. K, Pham Q.V, Gadekallu T. R, Krishnan S, S. R. Chowdhary, C. L. Jalil Piran M, 2020, “Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey”, Sustainable Cities and Society, 102589. doi:10.1016/j.scs.2020.102589

[12] Mehta P, Sandfort V, Gheysens D, Braeckevelt G.J, Berte J & Summers R. M, 2019, “Segmenting the Kidney on CT Scans Via Crowdsourcing”, 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). doi:10.1109/isbi.2019.8759240

[13] Fatma MuberraYener, AyseBetulOktay. (2020), “Diagnosis of Covid-19 with a Deep Learning Approach on Chest CT Slices”, IEEE Access, 1–1.10.1109/TIPTEKNO50054.2020.9299266

[14] Naufal Hilmizen, AlhadiBustamam, AlhadiBustamam, (2020), “The Multimodal Deep Learning for Diagnosing Covid-19 Pneumonia from Chest CT-Scan and X-Ray Images”, 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), doi: 10.1109/ISRITI51436.2020.9315478

[15] Shadmi R., Mazo V., Bregman-Amitai O., &Elnekave, E, (2018), “Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest CT”, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). doi:10.1109/isbi.2018.8363515

[16] El-kenawy, E.-S. M., Ibrahim, A., Mirjalili, S., Eid, M. M., & Hussein, S. E. (2020), “Novel Feature Selection and Voting Classifier Algorithms for COVID-19 Classification in CT Images”, IEEE Access, 1–1. doi:10.1109/access.2020.3028012

[17] Jakubicek, R., Chmelik, J., Ourednicek, P., & Jan, J. (2019)., “Deep-learning-based fully automatic spine centerline detection in CT data”, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2019.8856528

[18] Bhandary A, Prabhu G. A, Rajinikanth V, Thanaraj K. P, Satapathy S. C, Robbins D. E, Raja N. S. M, 2019, “Deep-Learning Framework to Detect Lung Abnormality – A study with Chest X-Ray and Lung CT Scan Images”, Pattern Recognition Letters. doi:10.1016/j.patrec.2019.11.013

[19] Haritha .D, Swaroop. N, Mounika. M, (2020), “Prediction of Covid-19 Cases using CNN with X-rays”,IEEE Access, 1–1. doi:10.1109/access.2020.30203416

[20] Yongzhao Xu, Gabriel Holanda, Lu´ısFabr´ıcio. de F. Souza, Hercules Silva, Adriell Gomes, Iagson Silva, Marcos Ferreira Junior, Chuanyu Jia, Tao Han, Victor Hugo C. de Albuquerque & Pedro P. Rebouc¸as Filho, “Deep Learning-Enhanced Internet of Medical Things to Analyze Brain CT Scans of Hemorrhagic Stroke Patients: A New Approach”, (2020), DOI 10.1109/JSEN.2020.3032897, IEEE Sensors Journal

[21] Wang, Jun; Bao, Yiming; Wen, Yaofeng; Lu, Hongbing; Luo, Hu; Xiang, Yunfei; Li, Xiaoming; Liu, Chen; Qian, Dahong (2020), “ Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images”, IEEE Transactions on Medical Imaging, (), 1–1. doi:10.1109/TMI.2020.2994908

[22]SertanSerte, Hasan Demirel, 2021, “Deep Learning for Diagnosis of Covid-19 using 3D CT Scans”, Computers in Biology and Medicine, doi: 10.1016/j.compbiomed.2012.104306

[23] Mohammad Rahimzadeh, Abolfazl Attar, Seyed Mohammad Sakhaei, 2021, “A Fully Automated Deep Learning-Based Network for Detecting COVID-19 From A New and Large Lung CT Scan Dataset”, Biomedical Signal Processing and Control, doi: 10.1016/j.bspc.2021.102588

[24] Kuruvilla J &Gunavathi K, 2014, “Lung cancer classification using neural networks for CT images”, Computer Methods and Programs in Biomedicine, 113(1), 202–209. doi:10.1016/j.cmpb.2013.10.011

[25] Jin, L., Yang, J., Kuang, K., Ni, B., Gao, Y., Sun, Y., … Li, M. (2020), “Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet”, EBioMedicine, 62, 103106. doi:10.1016/j.ebiom.2020.103106

[26] Tanvir Mahmud, Md. JahinAlam, Sakib Chowdhury, Shams Nafisa Ali, Md Maisoon Rahman, Shaikh Anowarul Fattah, (2020) “CovTANet: A Hybrid Tri-level Attention Based Network for Lesion Segmentation, Diagnosis, and Severity Prediction of COVID-19 Chest CT Scans”,10.1109/TII.2020.3048391

[27] Himmah F, Sigit .R&Harsono T, (2018), “Segmentation of Liver using Abdominal CT Scan to Detection Liver Disease Area”, 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC). doi:10.1109/kcic.2018.8628561

[28] Hidayatullah R. R, Sigit R &Wasista S, 2017, “Segmentation of head CT-scan to calculate percentage of brain haemorrhage volume”, 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), doi:10.1109/kcic.2017.8228603

[29] Nurhayati, O. D., &Windasari, I. P, 2015, “Stroke identification system on the mobile based CT scan image”, 2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). doi:10.1109/icitacee.2015.7437781

[30] Kido, S., Hirano, Y., & Hashimoto, N. (2018). “Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN)”. 2018 International Workshop on Advanced Image Technology (IWAIT). doi:10.1109/iwait.2018.8369798.