

A CNN Based Method for Detecting Covid-19 from CT Images

Buket KAYA¹ , Muhammed ÖNAL² 

¹Department of Electronics and Automation, Firat University, Elazığ, Turkey

²Department of Ecoinformatic, Firat University, Elazığ, Turkey

(bkaya@firat.edu.tr, muhammedonal@hotmail.com.tr)

Received: Sep.3, 2021

Accepted: Sep.16, 2021

Published: Oct.20, 2021

Abstract— COVID-19 outbreak first emerged on December 31, 2019 in Wuhan, China. The Novel Coronavirus Disease is caused by the SAR-CoV-2 virus, which causes respiratory symptoms such as fever, cough, and shortness of breath. While scientists continue their fight against

SARS-CoV-2 (2019-nCoV), one of the deadliest viruses in the last century, with tests to help diagnosis and prognosis, drug and vaccine discovery, Information Technologies mostly continues to work on early diagnosis, prognosis and prediction. The aim is to reveal systems with low margin of error that will alleviate the workload of healthcare professionals, as well as early diagnosis and initiation of treatment. Deep Learning and Computer vision is the most commonly used. Two class (covid, non-covid) classification solution, using the Artificial Intelligence Techniques, have been examined in this paper. CNN architecture, has been created to develop an model to disease detection process of COVID-19(2019-nCoV) virus infected patients from CT images consisting of NON-COVID and COVID classes. We have proposed the classifying of CT images using the 2 Convolutions and pool layers with the model which shortening the time for classification and achieved an accuracy of nearly 94.21%. Results show that the used model attains provide highly satisfying results and can be used for any image classification.

Keywords: *Deep Learning, CNN, 2019-nCoV, SARS-CoV-2, COVID, Classification, Chest CT images.*

1. Introduction

COVID-19 (SARS-CoV-2), outbreak first emerged from Wuhan China, has spread worldwide as a global pandemic. As of 4 Aug 21, more than 200 millions of people have been infected including 4.2 million deaths and 3.9 billion vaccine doses worldwide due to the pandemic [1]. COVID-19 (SARS-CoV-2) is one of the largest events in human history which has affected in such proportion globally in last century. The pandemic has not only affected people's routine activities, but also caused mental and psychological stress as well as significant financial losses, economic recession and health obligations [2,3]. The slowdown of the epidemic as a result of the vaccination studies carried out by some countries recently gave hope. In order to prevent the potential contagiousness of virus-infected patients, radiologists use CT images, which is one of the most effective technique to detect COVID-19 infection in patients. Because, as all health institutions have CT imaging devices, thus, CT based COVID-19 classification can be implemented to determine COVID-19 infected patients at a quickly. When radiologists evaluate CT images, specific and consistent features have been identified in nearly all COVID-19 patients. These, one ground glass opacities in the early stage and the other one is pulmonary embolism demonstrating linear consolidation in the latter stages, have been identified as prominent signs in detecting the infection of virus [4]. In order to alleviate the increasing health burden in epidemics and

to take the epidemic under control, this manual image evaluation method takes a lot of time. Therefore, As an alternative is to find a system that can detect and classify early and definitively to prevent the spread of COVID-19 disease. For this reason, Information Technologies continues its studies mainly on early diagnosis, prognosis and prediction.

ML(Machine learning), DL(deep learning) and AI based approaches have been used for detection and classifications various diseases [5]. For this reason, CNN, which is one of the DL(deep learning) architectures used in the field of medical imaging, has been very effective and many studies have developed with successful results. Our study aimed to devise a DL based model to classify the CT images belonging to 2 classes. Non-COVID and COVID infected person. For collecting the dataset, open-sourced online sources were referred. These dataset has granted us to train Neural Networks and provide highly getting results.

In this study, a disease detection of classifying model is implemented and the main contributions are as follows:

1. The dataset has been prepared with two-class of CT images (covid and non-covid) from open-sourced online sources were referred.
2. Performance improvement methods have been applied in the model to make better the performance of model.
3. Some normalizations were made to reduce the overfitting and thus the generalization error in model.

2. Dataset

In this paper, accepted databases, including Google Scholar, Kaggle, Github, Science Direct and Radiopaedia have been used to search for 2019-nCoV papers, CT images and other datasets. The papers, images and other datasets have been search using the keywords COVID, 2019-nCoV, DL, CNN, DL Architecture, CT images, Image Classification, Optimizers and Accuracy.

In this paper, one datasets from Kaggle including 2481 CT images are used. The data in this dataset were collected from real patients in hospitals in Sao Paulo, Brazil. The public available data set coming from Refs. [<https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset>] which includes 1252 COVID CT images and 1229 non-COVID CT images. The dataset was pre-processed and stored in common image format (PNG). The total dataset is used in 80% for training of the neural network and 20% for testing of the neural network. The information about dataset are summarized in Table I and Table II.

Table1. Dataset

	<i>NON COVID CT Images</i>	<i>COVID CT Images</i>	<i>TOTAL CT Images</i>
Num. of CT Images	1229	1252	2481

Table 2. Training and Test Rates

	<i>Name of Classes</i>	<i>Training(80 %)</i>	<i>Test(20 %)</i>	<i>Total</i>
1	non COVID	984	245	1229
2	COVID	1001	251	1252

Randomly selected samples of COVID and NON-COVID CT images with different sizes are shown below. (Fig.1) Image pre-processing is required.

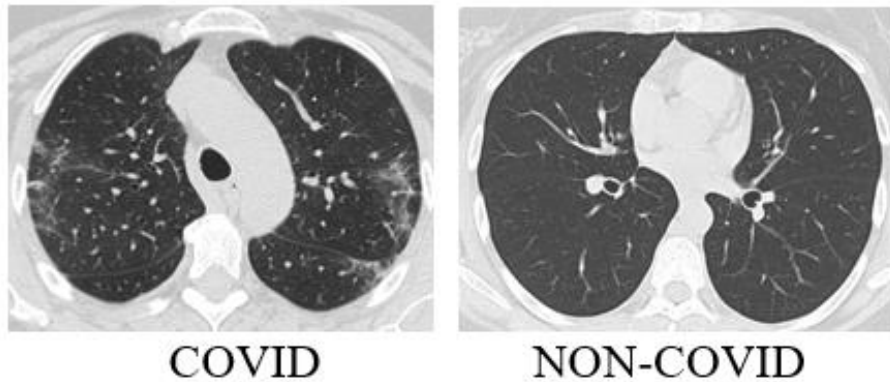


Figure 1. CT image samples of two classes as normal (non-covid) and covid

Images are not at the same size. On the other hand, images are converted to the same size for training.

3. Material and Method

Diagnosing and detecting covid from medical images such as pneumonia images is basically an image classification problem. In this type of classification, since it is important to capture the features of the image, some problems are frequently experienced in image classification. Images contain different textural features, image noise, time cost, etc. problems negatively affect the successful classification of the image. In our solution, DL(Deep Learning), which is one of the ML (machine learning) techniques, is used for covid diagnosis from chest CT images. In this context, CNN (Convolutional Neural Network) based on ANN (Artificial Neural Network), which is one of the DL architectures, is used to classify CT images (covid, non covid). In many studies, it has been seen that CNN is superior to traditional machine learning methods with its classification performance [6].

This section presents all the details of the suggested model and the steps taken in the working process. First of all, the DL method used in the study are briefly mentioned, then the processes of obtaining the data set and finally, the process of creating the model is included. Python open source software and Keras Library has been used to build the model. CNN algorithm was used in the study to classifying CT images in the data set.

CNN techniques are DL approach. DL is a method that provides solutions for nonlinear problems using more than one layer. It is generally used in areas such as video analysis, image classification, speech recognition information retrieval, object recognition and natural language learning [7-8]. However, when making a Deep Learning application, a few questions should be considered in determining which model to choose for a particular problem;

1. Should be a generative or discriminative model used?
2. What are the properties of data?
3. what are size of the input values?

In this sense, if the purpose is classification, it would be appropriate to use a distinctive model.

3.1. Deep Learning (DL)

DL (Deep learning) uses many layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input. Thus, DL(Deep Learning) is a new ML(machine learning) technique that can automatically extract features from the raw dataset. Essentially, we can say that deep learning is a more advanced version of traditional machine learning. There are two types of deep learning, supervised and unsupervised.[9]

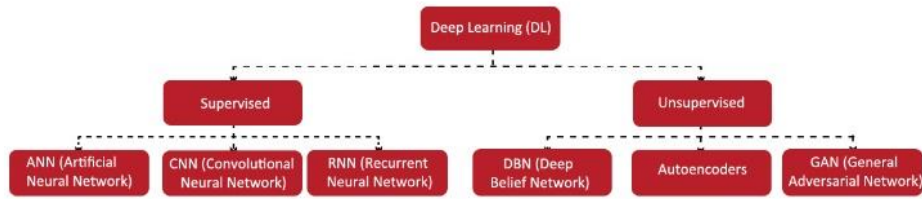


Figure 2. Types of Deep Learning

3.2. Convolutional Neural Network (CNN)

It is one of the most successful models of deep learning. Convolutional Neural Network is also known as CNN or ConvNet. CNN, in its simplest definition, is a deep learning model that can take an input image and distinguish between various objects in the image. Convolutional Neural Network has multiple layers. The first few stages of a basic CNNs consist of Convolution Layer and Pooling Layer. In the last stage, there is a fully connected layer. Then there is the classification layer [10].

When we examine the subject of Deep Learning and computer vision, the first model we will come across CNN. The biggest feature that distinguishes this type of networks from other Deep Learning methods is that it takes the image itself as input. In other DL models, the input is in vector form. CNN uses 2D and 3D images as direct input data.

The most important advantages of CNN are as follows:

1. By using convolutional filter and pooling layers, we can extract more information from the image and significantly reduce the network parameter to be trained.
2. Another important feature of CNN is *weight sharing*. (partial addition)
3. We can talk about the *local connectivity* feature of CNN (partial independence) (Fig.3) [11].

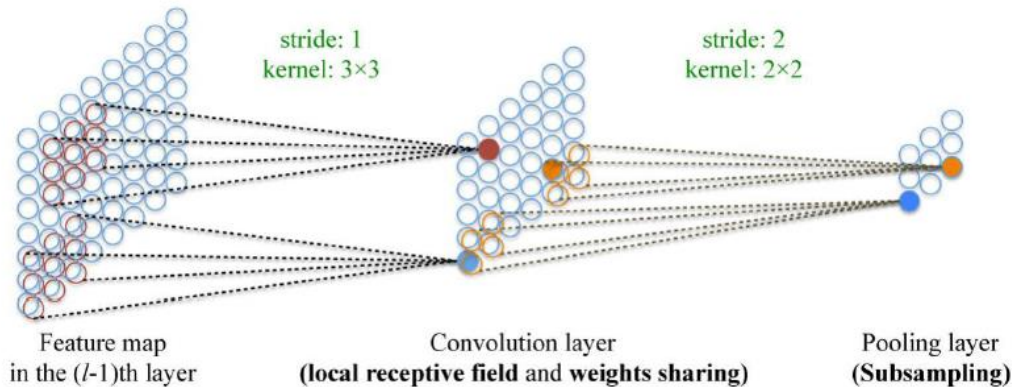


Figure 3. Three important mechanisms of convolutional networks (local connectivity, weight sharing and size reduction) [12]

3.3. Elements of CNN

CNN is a widely used method in tasks such as image classification and recognition today. But understanding exactly how these networks work can be difficult at times. In this section, we will talk about what elements consist of CNN and how it works.

A CNN architecture processes data in four stages. These are respectively;

- Convolution

- Non-Linear
- Pooling
- Fully Connected Layer

A sample CNN network that contains all these quantities are as shown in the figure. (Fig.4)

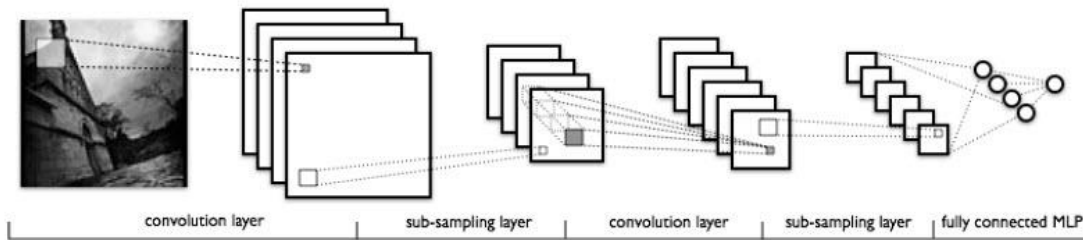


Figure 4. A CNN example [13]

As seen in Figure 5, CNN consists of multiple trainable sections placed one after the other, followed by a training classifier. After CNN receives the input data, the training process is completed by actualize layer-by-layer operations. Finally, it gives a result output for make comparison with correct result. An error occurs as much as the difference between the result (found result) and the desired result. Back propagation (backward propagation of errors) algorithm is used to transfer this error to all weights. With each iteration, the weights are revised to decrease the error.

Input Layer, one of the layers that make up the CNNs; It forms the first layer of the Convolutional Neural Network and in this layer, the data is given to the network as raw data. One of the most important features of this layer is that the size of the data in the layer is very important for the model to be designed be successful. Because, the higher the size of the input image, high performance is achieved with the success of the network.

The next layer is the **Convolution Layer**; This layer used to determine features. Combines the input and passes the result to the next layer. This layer can also be called the transformation layer. The reason why it is called a transform layer is because the transform operation circulates a particular filter over the entire image. These filters can be in different sizes. (2x2, 3x3, 5x5 e.g). These, create the output data by implementing convolution to the images from the former layer. As a result of this process (convolution process), feature maps are occur. The feature map is the regions where features specific to each filter are explored. After the convolution layer comes the flattened linear unit layer called **ReLU**. The outputs of the Convolutional Neural Network neurons are the rectifier unit that is activated. It is also known as the activation layer. This layer pulls negative values to zero with the effect it has on the input data. Since certain mathematical operations are applied in the convolution layer, the network is in a linear structure and the ReLu layer is applied to convert it to a nonlinear structure. The use of this layer provides faster learning of the network. After this layer, the **pooling layer** usually comes. The main purpose of the pooling layer is to decrease the input size for the next convolutional layer. The decrease in size causes loss of information (*loss of information*). Information loss, on the other hand, occurs less computational load for the next network layers, and also turns it into an advantage as it prevents the system from *overfitting*. Pooling layer is not mandatory in Convolutional Neural Network. Some models do not use this layer. Like the convolution operation in the first step, specific filters are defined in the pooling layer. While *MaxPooling* is generally preferred because it performs better, it performs this process by taking the maximum values of the pixels in the image by moving the filters on the image according to a specific stepping value (*stride*). Since the specified stride value can be increased or decreased, the output size also changes according to the realized stride value. When we examine the CNN architecture in Figure 5, we see that the **fully connected layer** comes after the successive convolution and **ReLU** and **pooling layers**. However, before the **fully connected layer**, there is the **Flatten Layer**. In the most basic terms, the mission of this layer is to make ready the data at the input of the fully connected layer. Neural

networks take input data from a one-dimensional array. This data, one-dimensional array of matrices from the convolution and pooling layer. We can say that the fully connected layer is a standard layer used in classification. This layer is connected to all neurons of the previous layer. When we train a model with big data in a convolutional neural network, sometimes things may not go as we want. One of the reasons for this is undoubtedly that the network makes memorization (*overfitting*). For this reason, a layer called **Dropout** is used so that the network does not memorize. This is accomplished by removing some nodes of the network while doing this. When designing a model, this number can be reduced starting from Dropout, 0.5, until the model reaches maximum performance. (*Fine Tuning*). **Dropout** can only be used in the Fully Connected Layer in the CNN. **Batch Normalization**, on the other hand, makes the CNN network more organized and ensures that it is stable/optimized during training. The **Classification Layer** is a layer that latter the Fully Connected Layer and is classified in this layer of the model. The output value of this layer is equal to the count of objects to be classified. Since the objects in the classification give a certain value in the range of 0-1, the output that produces a result close to 1 will be understood as the predicted object by the network.

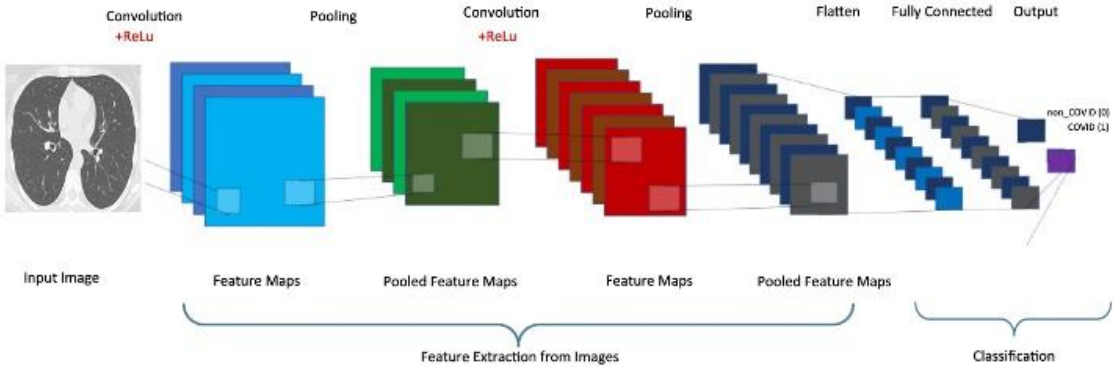


Figure 5. An example classifier and CNN architecture of the application.

3.4. Architecture of Model

The architecture of our model consists of several important steps to classify whether or not covid pneumonia is present from chest CT images. These steps are tried to be shown in the Block Diagram given in Figure 6. First, the data required for training and validation of the model are collected in accordance with the modality. The collected data is preprocess and shuffled, resized, normalized to maintain for stability. After this stage, it is categorized according to the classification of the model. Using deep learning networks, the model is trained and validated. Finally, the performance of the model is appraised and analyzed with some important metrics.

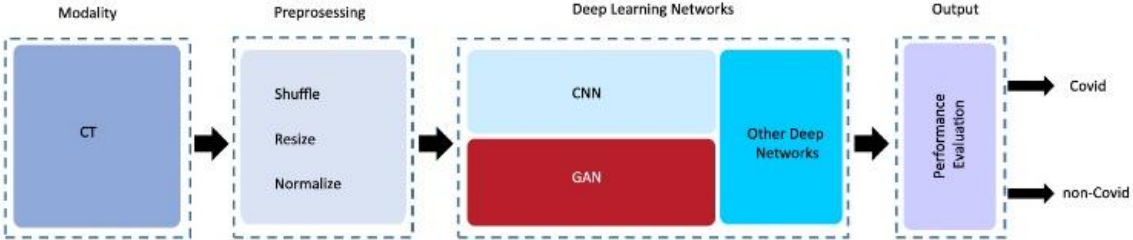


Figure 6. Block diagram of the model

As seen in Figure 7, the model we developed for the solution of the problem consists of 19 layers. Our model accepts the input image as 150x150 and transmits it to the convolution layer. The Kernel size of the conv. layers in the second, sixth and tenth layers is equal to 3x3. After each conv. layer, a pooling layer process is applied to extract feature maps from the image. However, before each pooling layer an activation layer (ReLU) and then make our network more organized and optimized during training were used a BatchNormalization layer. The main reason for using the pooling layer in our model is that it reduces the input size for the next convolution layer, as stated in its definition. We set the Dropout Rate to $p = 0.5$, which provides for max. regulation. Otherwise, if all the weights are learned together in the networks, normally some links will gain more predictive ability than others. In such a case, as the network is trained iteratively, these strong connections are learned more while weaker ones are ignored [14]. In order to avoid overfitting, our model also has used Dropout also in Fully Connected Layer in addition to the BatchNormalization layer we used after each convolution layer. Sigmoid activation function, which is very effective in binary classification, is used in the classifier layer of our model. In the output layer, our result has classified as 0 or 1.

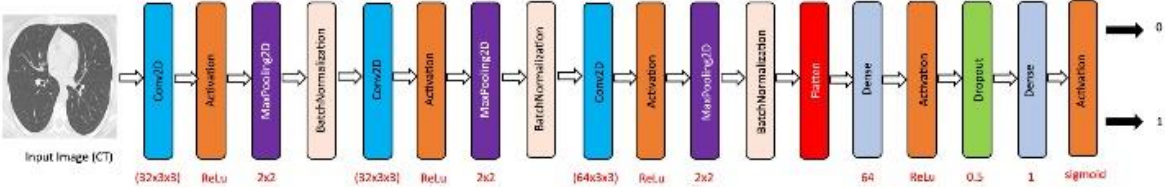


Figure 7. Architecture of model

The basic parameters we used in training our model are given in Table 3.

Table 3. Training Parameters

<i>Prog.Lang and Lib.</i>	<i>Optimization</i>	<i>Epoch</i>	<i>Loss</i>	<i>Learning Rate</i>	<i>Batch Size</i>
Python - Keras	adam	50	Binar y-cross entro py	0.01	64

Then, the training process that the network (CNN) will perform will be as below:

1. The parameters (weights) that all filters have are given random initial values.
2. The CNN takes a training image as input, goes through the forward propagation step, producing a possibility value for each class as a result.
3. The total error is computed (obtained by summing the values produced for all classes).
4. With the back propagation algorithm, the derivatives of the error according to the weights are calculated and the weights and other parameters are updated according to the optimization algorithm used.
5. Repeat steps 2-4 for all images. [15]

4. Evaluation Metrics

Major evaluation metrics commonly used to measure the quality of classification models were used to evaluate our model. The selected metrics are as follows.

- Confusion Matrix
- Sensitivity (Recall)
- Specificity
- Precision
- F1-Score
- Accuracy

The calculated parameters of these metrics are based on the True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) ratios as shown in Fig 8.

Performans Metrikleri	Formül
Sensitivity	$TP / (TP + FN)$
Specificity	$TN / (TN + FP)$
Precision	$TP / (TP + FP)$
F-Score	$2TP / (2TP + FP + FN)$
Accuracy	$(TP + TN) / (TP + FN + FP + TN)$
MCC	$\frac{(TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$

Figure 8. Performance Metrics of Model

Confusion Matrix is a performance measure for the classification problem. It contains four different combinations of estimated and actual values. (Fig. 9)

		PREDICTIVE VALUES	
		POSITIVE (1)	NEGATIVE (0)
ACTUAL VALUES	POSITIVE (1)	TP	FN
	NEGATIVE (0)	FP	TN

Figure 9. Confusion Matrix

5. Results

The results obtained using the Python Keras library are based on the last cycle, not the total. ReLU function is used in layers (except output).

The CNN-based model, whose details are shared in the "Architecture of Model" section, was used in the classification of Chest CT image sections with Covid Pneumonia symptoms and normal findings. As a result of 50 epochs, each of which lasted an average of 140 seconds, a success rate of 94.21% was achieved. The graphic representation of the numerical values of the training phase is given in Figure 10.

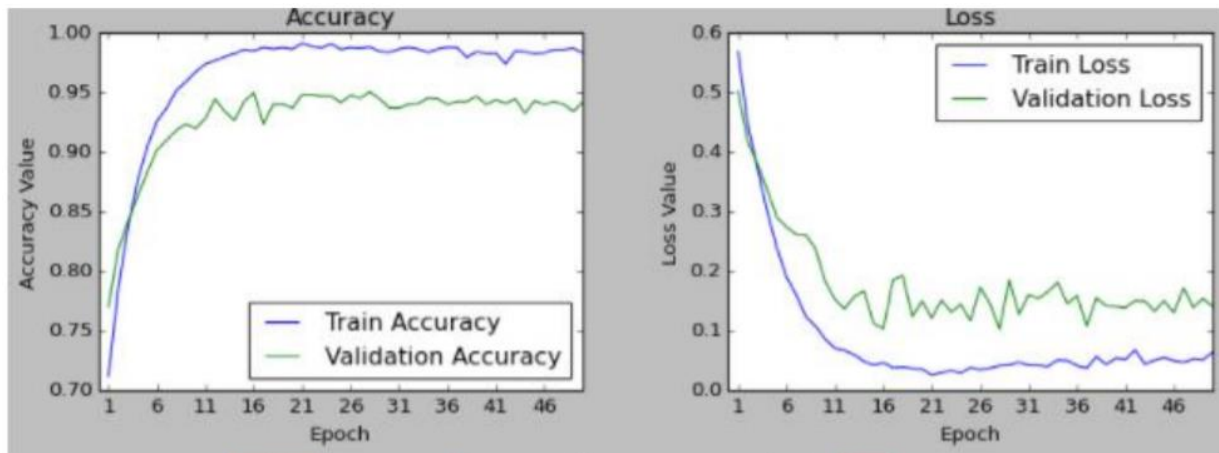


Figure 10. Training Stage Success and Loss Graphs

The proposed model (Fig.10) used at this stage of our study was trained for 50 epochs. 2481 CT images were used two parts as 80% for training and 20% testing.

The outputs of the evaluation metrics used in the testing phase of our model are given in Figure 11.

	precision	recall	f1-score
yok	0.95	0.97	0.96
var	0.94	0.91	0.92
accuracy			0.95
macro avg	0.94	0.94	0.94
weighted avg	0.95	0.95	0.95

Figure 11. Evaluation Metric Outputs

6. Conclusion

The performance of the suggested model is directly proportional to the number of data. Using more data has a positive effect on model accuracy and performance. The findings reveal that the proposed model can be easily used in the healthcare for any classification problem.

Although the CNN method is quite successful in extracting the qualitative features of the data and summarizing these features through convolutional operations, it basically has three important problems. Due to these problems, the CNN model cannot convey some important features and data loss occurs. These:

1. Translational invariance
2. Viewpoint invariance
3. Data loss due to max-pool [16]

In the next phase of the study, different algorithms or techniques can also be used to achieve higher success rates. In addition, since the subject of the study is health, more accurate comments can be obtained about the results obtained by taking the opinions of the experts in this field. As a result, the proposed model's ability to diagnose or detect in a short time, with a low error rate and at minimum cost, reveals the potential of artificial intelligence techniques to be used in the field of health.

References

- [1] <https://covid19.who.int/>, last accessed on 9 Jun 21.
- [2] Rajkumar RP. COVID-19 and mental health: a review of the existing literature. *Asian J Psychiatry* 2020;52 . <https://doi.org/10.1016/j.ajp.2020.102066>
- [3] Luo M, Guo L, Yu M, Jiang W, Wang H. The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public – a systematic review and meta-analysis. *Psychiatry Res* 2020;291 . <https://doi.org/10.1016/j.psychres.2020.113190>
- [4] URL:<https://radiologyassistant.nl/chest/covid-19/covid19-imaging-findings>, last accessed on 19 Oct 20.
- [5] Fraiwan L, Hassanin O, Fraiwan M, Khassawneh B, Ibnian AM, Alkhodari M. Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers. *Biocybern Biomed Eng* 2021;41(1):1–14. <https://doi.org/10.1016/j.bbe.2020.11.003>.
- [6] Chen, X., Xiang, S., Liu, C. L., & Pan, C. H., Vehicle detection in satellite images by hybrid deep convolutional neural networks, *IEEE Geoscience and remote sensing letters*, 11 (10), 1797-1801, 2014.
- [7] Özkan, İ. & Ülker, E., “Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri”, *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104, 2017.
- [8] Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu Perez, J.,Lo,B., & Yang, G.Z., “Deep learning for health informatics”. *IEEE Journal of Biomedical and health informatics*, 21(1), 4-21, 2016.
- [9] L. Deng and D. Yu, “Deep Learning: Methods and Applications,” *Found. Trends® Signal Process.*, vol. 7, no. 3–4, pp. 197–387, 2014.
- [10] Nurettin Öner, İsmail Ağırbaş, *Bilgisayar Tomografisi ve Manyetik Rezonans Görüntülemenin Sağlıkta Teknoloji Değerlendirme ve Maliyet-Fayda Analizi ile Değerlendirilmesi*, *Sağ. Perf. Kal. Derg* (8), 147-163, 2014
- [11] Zhou, S. K., Greenspan, H., and Shen, D., "Deep Learning for Medical Image Analysis", 1. Ed., Academic Press, London; San Diego, 458 (2017).
- [12] Shen, D., Wu, G., and Suk, H.-I., "Deep learning in medical image analysis", *Annual Review Of Biomedical Engineering*, 19 (1): 221–248 (2017).
- [13] Internet: Deep Learning, "What is the effect of adjusting BatchSize on the training effect?" <https://www.programmersought.com/article/2840518935/>
- [14] Frid-Adar,M.,Diamant, I.,Klang E.,Amitai,M.,Goldberger J.,Greenspan H.,GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification.*Neurocomputing*. 321-331,2018.
- [15] Internet: ujjwalkarn, "An Intuitive Explanation of Convolutional Neural Networks", <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/> (2018).
- [16] Sabour, S., Frosst, N., and Hinton, G. E., "Dynamic routing between capsules", *ArXiv:1710.09829 [Cs]*, (2017).