2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)

COVID-19 Detection Using Deep Learning Methods

Mehmet Sevi
IT Department
Muş Alparslan University
Muş, Turkey
m.sevi@alparslan.edu.tr

İlhan AYDIN
Department of Computer Engineering
Fırat University
Elazığ, Turkey
iaydın@fırat.edu.tr

2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI) | 978-1-7281-9675-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICDABI51230.2020.9325626

Abstract*—According to the world health organization, the coronavirus epidemic threatens the world's health system every day. Health resources in most countries are either insufficient or not fairly shared. There are various problems such as the number of health personnel, the number of beds, or the number of intensive care units. Using limited resources at the optimum level is the key to the country's health systems to overcome this epidemic. Disease detection is an important factor in preventing the epidemic. The higher the success, the more controlled the spread of the virus. Whether the person has a virus or not is usually done by the PCR test. In addition to the PCR method, chest x-ray images can be classified with deep learning methods. Deep learning methods have become popular in academic studies by processing multi-layered images in one go and by defining manually entered parameters in machine learning. This popularity reflected positively on limited health datasets. In this study, it was aimed to detect the disease of people whose x-rays were taken for suspected COVID-19. In such COVID-19 studies, a binary classification has generally been made. The data set includes chest x-rays of patients with COVID-19, viral pneumonia, and healthy patients. Before the classification process, the data augmentation method was applied to the data set. These three groups have been classified through multi-class classification deep learning models.

***Index Terms*—deep learning, COVID-19, viral pneumonia, multi-class classification, data augmentation**

I. INTRODUCTION

In this study, it is aimed to classification healthy patients, COVID-19, and viral pneumonia cases. Deep learning is a machine learning method. It allows us to train artificial intelligence to predict outputs with a given data set. Both supervised and unsupervised learning can be used to train artificial intelligence. Deep learning is used in voice and face recognition, disease detection, defense, and security areas. The word deep in deep learning represents artificial neural networks. Artificial neural networks are inspired by the human brain. Just like the human brain, it consists of neurons. The difference between them is the amount and speed of learning. In other words, data set and processing power are needed to train artificial neural networks.

The quality of machine learning methods depends on choosing the right features [1]. Various preprocesses, size reduction, feature selection, etc. transactions are made. In order to reduce the cost at this stage, it is necessary to get rid of the dependence on features [2]. This is where deep learning comes into play. Deep learning takes care of these things we do in machine learning. Deep learning uses many nonlinear layers for feature extraction and feature modification. In sequential layers, the exit of the previous one is the entrance of the next. Deep learning makes a hierarchical selection that best represents the data, rather than manual feature selection.

The data set used in the study was created by a few researchers from Qatar and Dhaka universities. The data set includes COVID-19 positive cases, healthy patients, and chest radiographs of viral pneumonia patients [3].

In this study, it is aimed to solve the multi-class classification problem. So, there are more than two classes to predict. To explain, there are three patient diagnoses. This is an important type of problem that can be practiced with neural networks because three class values require special processing.

In another study in this area, Rustam et al. tried to predict the future numbers of Covid-19 patients [4]. In addition to the number of infected patients, they also tried to estimate the number of patients who died, and the number of patients recovered over the next 10 days. As a result of the research, exponential smoothing gave the best result among all the models used. Linear regression, least absolute shrinkage, and selection operator gave the best result after exponential smoothing. Support vector machine (SVM) gave the worst accuracy rate. In the study, country, city, longitude, latitude, number of deaths, number of patients recovered, and the number of cases were used for estimation.

In another study, Roy et al. investigate deep learning-based solutions for lung diseases [5]. In the study, it is presented a dataset of LUS images collected from several Italian hospitals. The dataset contains labels with video and pixel levels. The experiments in the data set used showed satisfactory results. The study has shown that COVID-19 can be diagnosed from the LUS data. The study has also guided deep learning studies on COVID-19 in the future. U-net ++ gave the most successful result with a rate of 97% [5].

A study by Sethi et al. made emotion detection in Tweets containing COVID-19 in the literature [6]. The data set used in the study was created automatically through the Twitter API. Three different data sets were used in the study, and binary, multiple, and cross-classification were made. The binary classification has the most successful results among these three classification tasks. Tweets are classified as positive and negative in the binary classification process. Machine learning algorithms Decision Tree and SVM achieved about 90% success [6]. In multiple classifications, tweets are classified as neutral, positive, and negative. In the multiple classification process, the same machine learning algorithms achieved less success than the binary classification process. They replicated this performance in different data sets. The cross-dataset classification process is the most unsuccessful method in the article.

In another tweet-based study, Long et al. examined text-based tweets and suggested a system for governments and

municipalities to understand citizens during the epidemic [7]. Developed NeedFull system collected data from Twitter and analyzed the emotional needs of people during the pandemic process with machine learning methods created after the feature selection and normalization stages [7].

The model developed in the second part of the article will be discussed. CNN, VGG, VGG19, and InceptionV3 deep learning models were used in the article. The experimental results will be discussed in the last part.

II. Proposed Method

The chest X-rays of the patients contain information that determines whether the person is COVID-19. In this article, chest x-rays of the patients were classified as COVID-19, healthy and viral pneumonia. The data set includes chest X-ray images of 219 COVID-19 patients, 219 healthy patients and 219 viral pneumonia patients. This means a total of 657 images. Figure 1 shows sample chest X-ray images from the data set.

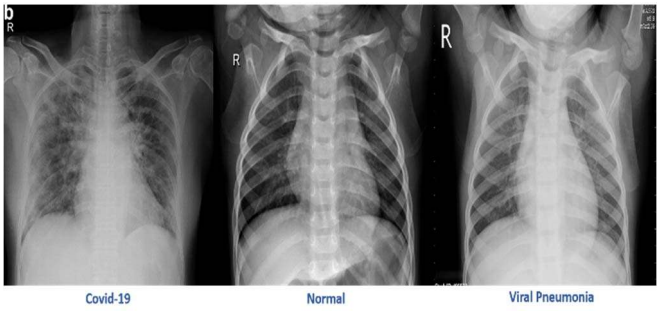

Fig. 1. Data set samples.

Confusion matrix is used to evaluate the performance of deep learning models. Confusion matrix is a matrix in which target predictions and actual values are compared to evaluate the classification performance used in deep learning algorithms. The criteria used in the evaluation of the classification results except the confusion matrix are as follows:

- Accuracy: It is the main evaluation element in the study. It is the ratio of the total number of correct predictions to the number of predictions tested [8].
- Precision: It is the ratio of the number of positive real value classified as positive value to the total of those classified as positive value [8].
- Recall: It is the ratio of the number of positive real value classified as positive value to all positive real value [8].
- F-Score: Its value shows the harmonic mean of the recall and precision values.

One of the ways to increase the accuracy rate in the study is by data augmentation. The over-fit problem of the deep learning model stands out in small data sets. Which means a low accuracy rate. We rotate the images in the data set by 15 degrees clockwise or counterclockwise to increase the accuracy. Thus, the data set has been expanded.

Deep learning algorithms have become more and more popular after the competition organized by Imagenet in 2012. It has started to be used more in academic research. Convolutional neural network (CNN) is one of the deep learning networks used for computer vision. The CNN algorithm was created by taking inspiration from the visual center of animals [9]. CNNs are structures designed to take images as input and are used effectively in computer vision. CNN consists of one or more convolutional layers and one or more fully connected layers, such as a standard multilayer neural network [9]. These; Convolution, Relu, Pooling, Flattening, and Fully Connected layers. The convolution layer is the main block of CNN [10]. The layer is responsible for perceiving the features of the images. Some filters are applied to the picture to extract the high and low-level features in the images. Filters usually contain multidimensional matrices and pixels. The feature map, which is the last matrix obtained as a result of the applied filters, is shaped according to the filters used. After the Convolution layer, the order is in the nonlinear layer. Since the activation process is done in this layer, it can be called the Relu layer. The relu activation function sets the negative values from the feature map to 0 [10]. The task of this layer is to reduce the sheer size of the representation and the number of parameters and calculations within the network. In this way, incompatibility in the network is checked. The main purpose here is to reduce the number of parameters we have and to keep the most critical parameters and to reduce the number of entries for the next layer. This reduces the computational cost for subsequent layers and prevents memorization. In the flattening layer, the main purpose is to prepare data for the last layer, Fully Connected layer. It converts matrices from Convolutional and Pooling layers into a one-dimensional array [10]. The last layer, the Fully Connected Layer, takes the data from the Flattening layer and performs the learning process through the Neural network. CNN scheme is given in Figure 2.

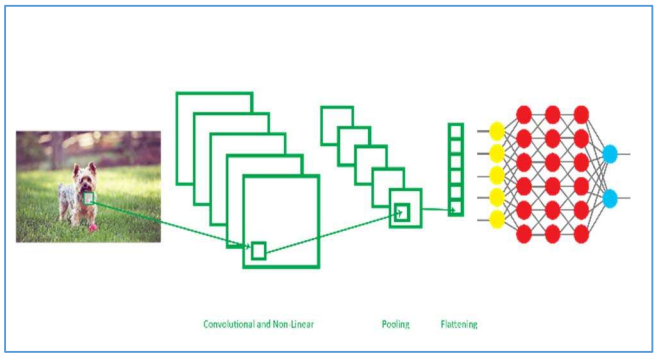

Fig. 2. CNN scheme [11].

As the name suggests, VGG-16 consists of 16 layers, and VGG-19 consists of 19 layers. Both of were the basis of the Visual Geometry Group (VGG). The group won the first and second places in the ImageNet Challenge 2014 [12]. Small filters (3x3) are used in convolution layers in VGG16 [13]. VGG16 consists of 13 convolution layers and 3 fully layers in Figure 3. There are 5 pooling layers of 2x2 size. The last layer contains the softmax layer. The input data coming with the softmax layer are classified. Relu is used as the activation function. VGG19 consists of 16 convolution layers and 3 fully connected layers in Figure 3. As seen in Figure 3, VGG19 has five pool layers like VGG16 and as the last layer, both have a softmax layer. VGG16 contains 138 million parameters, while VGG19 contains approximately 144 million parameters. The disadvantage of both VGG models is that they consume a lot of memory. VGG16 and VGG19 schemes are given in Figure 3.

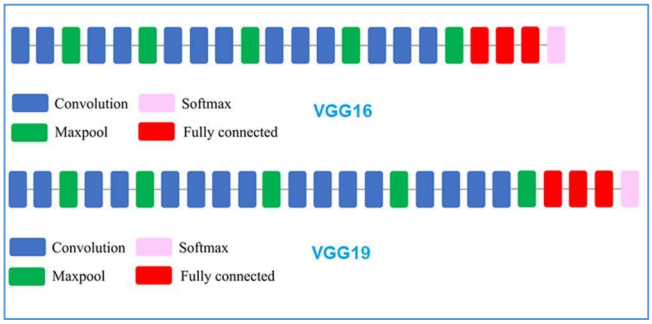

Fig. 3. VGG16 and VGG19 scheme [14].

InceptionV3 is a type of convolutional neural network model. It was developed by Christian Szegedy. It consists of multiple convolution and maximum pooling steps. It consists of a 42-layer deep neural network. [15]. In the last layer, it involves a fully connected neural network. It has 78.1% accuracy rate on ImageNet dataset. The model requires 299x299x3 image size [15]. The InceptionV3 scheme and layers of the model are given in Figure 4.

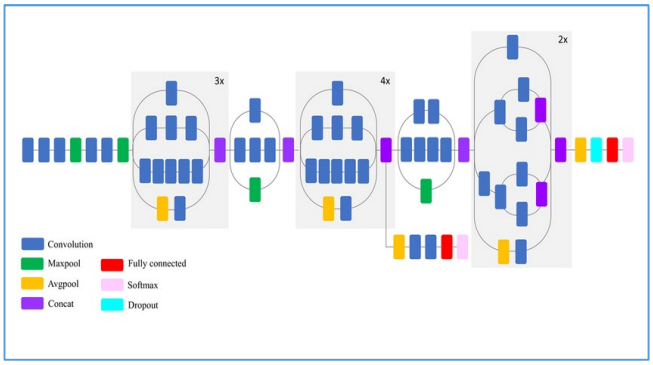

Fig. 4. IncepitonV3 scheme [16].

III. Experimental Results

Deep learning methods: CNN, VGG16, VGG19, and InceptionV3 are applied to the chest x-ray dataset. Tensorflow and Keras Python libraries were used in the article. In the 4 deep learning models used, Adam was used as optimizer, and Relu and softmax were used as activation functions.

Table 1 shows the confusion matrix and other evaluation criteria of the CNN model. The CNN model's accuracy rate is 85%. The closer the precision and recall values are to 1, the higher the model's ability to distinguish between classes.

TABLE I. Confusion Matrix of CNN

Predicted Class True Class	Covid	Normal	Viral	Precision	Recall	F1-Score
Covid	34	7	3	0.94	0.77	0.85
Normal	1	39	4	0.78	0.89	0.83
Viral	1	4	39	0.85	0.89	0.87

When val_loss (validation loss) starts to increase, val_acc (validation accuracy) starts to neither increase nor decrease in Figure 5. This means the model is not overfitting and the model is cramming values not learning. Overall accuracy is acceptable.

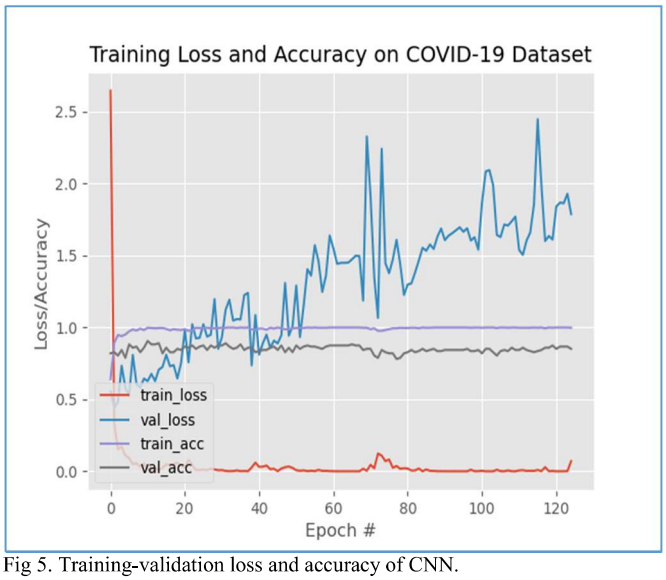

Fig 5. Training-validation loss and accuracy of CNN.

Table 2 shows the confusion matrix and other evaluation criteria of VGG16 model. The VGG16 model's accuracy rate is 93%.

TABLE II. Confusion Matrix of VGG16

Predicted Class True Class	Covid	Normal	Viral	Precision	Recall	F1-Score
Covid	42	0	2	0.91	0.95	0.93
Normal	0	43	1	0.98	0.98	0.98
Viral	4	1	39	0.93	0.89	0.91

When val_loss starts to decrease, val_acc starts to increase in Figure 6. This means the model is learning and working fine.

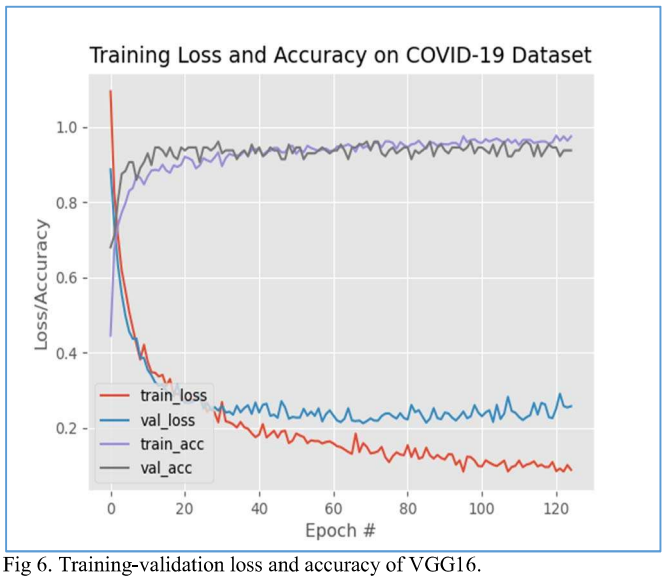

Fig 6. Training-validation loss and accuracy of VGG16.

Table 3 shows the confusion matrix and other evaluation criteria of VGG19 model. The VGG19 model's accuracy rate is 95%.

TABLE III. CONFUSION MATRIX OF VGG19

Predicted Class True Class	Covid	Normal	Viral	Precision	Recall	F1-Score
Covid	41	2	1	1.00	0.93	0.96
Normal	0	43	1	0.91	0.98	0.95
Viral	0	2	42	0.95	0.95	0.95

When val_loss starts to decrease, val_acc starts to increase in Figure 7 like that VGG19 loss and accuracy graphic. This means the model is learning and working fine.

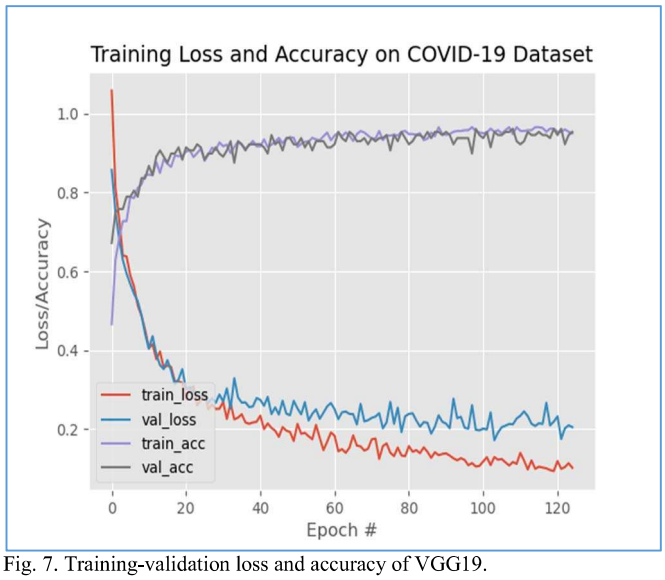

Fig. 7. Training-validation loss and accuracy of VGG19.

The accuracy rate of InceptionV3 is 48%. Table 4 shows the confusion matrix and other evaluation criteria of InceptionV3 model.

TABLE IV. CONFUSION MATRIX OF INCEPTIONV3

Predicted Class True Class	Covid	Normal	Viral	Precision	Recall	F1-Score
Covid	9	0	35	1.00	0.20	0.34
Normal	0	13	31	0.81	0.30	0.43
Viral	0	3	41	0.38	0.93	0.54

The overall accuracy in Figure 8 is not acceptable. InceptionV3 model has mistakes. These mistakes can disappear after which can change the model architecture or more data.

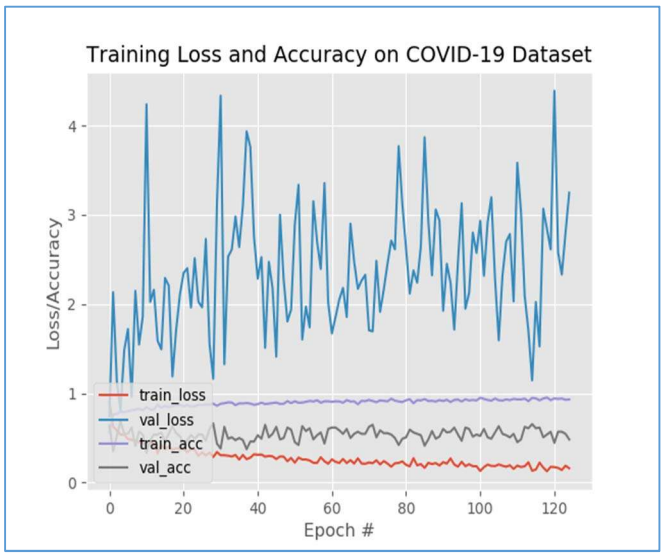

Fig. 8. Training-validation loss and accuracy of InceptionV3.

Figure 9 shows a comparison of the deep learning models.

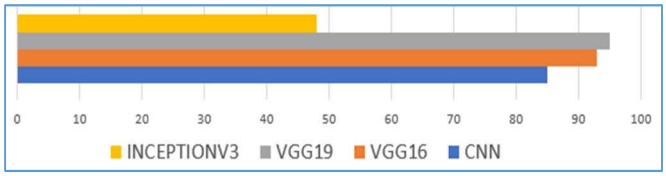

Fig. 9. Comparison of the deep learning models.

IV. CONCLUSİON

Up to now, a total of 657 chest X-ray images have been examined for the diagnosis of COVID-19 using deep learning methods. This number has been increased with the data augmentation technique. VGG19 is most the successful model that has 95% accuracy rate. COVID-19 patients, healthy patients, and viral pneumonia cases are classified successfully by VGG19 model. InceptionV3 is the most unsuccessful method for the dataset.

This article emphasizes the importance of the COVID-19 outbreak. When the number of x-ray images, ultrasound videos increases day by day the deep learning methods will more be used day by day in the health sector. The results of the studies in the literature show the importance of deep learning in combating the COVID-19 outbreak [17]. In future studies, the success ratio can be increased by strengthening the data set. Lung tomography can be used in addition to chest radiographs. By developing different deep learning models, success ratio and performance can be increased.

REFERENCES

[1] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," Artif. Intell., vol. 97, no. 1–2, pp. 245–271, Dec. 1997.

[2] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Trans. *Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.

[3] https://www.kaggle.com/tawsifurrahman/covid19-radiography-database COVID-19 Radiography Database. 1 July 2020.

[4] F. Rustam et al., "COVID-19 Future Forecasting Using Supervised Machine Learning Models," in IEEE Access, vol. 8, pp. 101489-101499, 2020, doi: 10.1109/ACCESS.2020.2997311.

[5] S. Roy et al., "Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound," in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2676-2687, Aug. 2020, doi: 10.1109/TMI.2020.2994459.

[6] M. Sethi, S. Pandey, P. Trar and P. Soni, "Sentiment Identification in COVID-19 Specific Tweets," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 509-516, doi: 10.1109/ICESC48915.2020.9155674.

[7] Z. Long, R. Alharthi and A. E. Saddik, "NeedFull – a Tweet Analysis Platform to Study Human Needs During the COVID-19 Pandemic in New York State," in IEEE Access, vol. 8, pp. 136046-136055, 2020, doi: 10.1109/ACCESS.2020.3011123.

[8] İ. Aydın, M. Sevi, U. M. Salur, "Makine Öğrenmesi Algoritmaları ile Sahte Twitter Hesapların Tespiti," in International Conference on Artificial Intelligence and Data Processing, 2018.

[9] G. Gündüz, İ. H. Cedimoğlu, "Derin Öğrenme Algoritmalarını Kullanarak Görüntüden Cinsiyet Tahmini," Sakarya Unıversıty Journal Of Computer And Informatıon Scıences, 2, 9-17, 2019.

[10] K. Shankar, V. Iyer, K. Iyer and A. Pandhare, "Intelligent Video Analytics (IVA) and Surveillance System using Machine Learning and Neural Networks," 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, pp. 623-627, doi: 10.1109/ICICT48043.2020.9112527.

[11] https://medium.com/@tuncerergin/convolutional-neural-network-convnet-yada-cnn-nedir-nasil-calisir-97a0f5d34cad/ Convolutional Neural Network (ConvNet yada CNN) nedir, nasıl çalışır? 15 Agust 2020.

[12] M. M. Leonardo, T. J. Carvalho, E. Rezende, R. Zucchi and F. A. Faria, "Deep Feature-Based Classifiers for Fruit Fly Identification (Diptera: Tephritidae)," 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, 2018, pp. 41-47, doi: 10.1109/SIBGRAPI.2018.00012.

[13] S. Toraman, B. Daş, "Evrişimsel Sinir Ağları Kullanılarak Normal Ve Göğüs Kanseri Hücreleri İçeren Genomların Sınıflandırılması" DÜMF Mühendislik Dergisi, 11-1, 81-90, 2020.

[14] https://www.researchgate.net/figure/Schematic-diagram-of-a-VGG16-and-b-VGG19-models_fig4_326421398/ Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery 15 Agust 2020.

[15] A. P. Rahmathunneesa and K. V. Ahammed Muneer, "Performance Analysis of Pre-trained Deep Learning Networks for Brain Tumor Categorization," 2019 9th International Conference on Advances in Computing and Communication (ICACC), Kochi, India, 2019, pp. 253-257, doi: 10.1109/ICACC48162.2019.8986151.

[16] https://www.researchgate.net/publication/326421398_Very_Deep_Convolutional_Neural_Networks_for_Complex_Land_Cover_Mapping_Using_Multispectral_Remote_Sensing_Imagery/ Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery 15 Agust 2020.

[17] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio and S. K. Antani, "Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays," in IEEE Access, vol. 8, pp. 115041-115050, 2020, doi: 10.1109/ACCESS.2020.3003810.