Covid-19 Classification Using Deep Learning in Chest X-Ray Images

Zehra KARHAN
Necmettin Erbakan University
Konya/TURKEY
zkarhan@erbakan.edu.tr

Fuat AKAL
Hacettepe University
Ankara/TURKEY
akal@hacettepe.edu.tr

Abstract— Covid-19 virus, which has emerged in the Republic of China in an undetermined cause, has affected the whole world quickly. It is important to detect positive cases early to prevent further spread of the outbreak. In the diagnostic phase, radiological images of the chest are determinative as well as the RT-PCR (Reverse Transcription-Polymerase Chain Reaction) test. It was classified with the ResNet50 model, which is a convolutional neural network architecture in Covid-19 detection using chest x-ray images. Chest X-Ray image analysis can be done and infected individuals can be identified thanks to artificial intelligence quickly. The experimental results are encouraging in terms of the use of computer-aided in the field of pathology. It can also be used in situations where the possibilities and RT-PCR tests are insufficient.

Keywords — Chest X-Ray Images, Coronavirus, Covid-19, Deep Learning, Transfer Learning

I. INTRODUCTION

The coronavirus (COVID-2019), quickly spread around the world and became a pandemic. It is crucial to detect covid-19 positive cases as early as possible to prevent further spread of this epidemic and treat affected patients rapidly [1-3].

RT-PCR (Reverse Transcription-Polymerase Chain Reaction) test is applied for the diagnosis of Covid19 [4]. The disadvantages are that the results of the RT-PCR test are long and some countries do not meet the need. With this disadvantage, case detection cannot be made early. Therefore, it increases the spread by the interaction of diseased and healthy individuals. At this stage, other methods (patient history, clinical symptom analysis, pathogenic test, computed tomography, etc.) are used for diagnosis. Diagnosis from chest images gains importance at the beginning of these methods [5-7].

Many studies have been conducted on the automating of the detection of the disease from chest x-ray images using artificial intelligence. Souza et al. Presented lung segmentation on chest x-ray images [6]. Lin et al conducted studies to identify cases of pneumonia with the purpose of early diagnosis [7]. Studies have accelerated on the diagnosis of chest X-ray images of Covid-19 suspect with the studies in this field. Deep learning, which has been used frequently recently, has been successfully applied to x-ray images in many areas. Abbas and et al. presented the DeTraC model that relies on a class decomposition to the classification of chest images [8]. Wang, Linda, and Wong proposed COVID-Net based on human-machine collaboration [9]. Ioannis et al have worked on Covid-19 detection using pre-trained convolutional neural networks on chest images [10].

In this study, chest X-ray images created using more than one data set were studied. Due to the small dataset, image augmentation techniques were used. Classification of Covid-19 cases was performed using the ResNet-50 model based on pre-trained convolutional neural networks on chest x-ray images. After 5-fold cross-validation, Covid-19 were identified. Performance classification, confusion matrix, and ROC curve are given.

The rest of this paper is organized as follows. Section 2 presents the materials and acquisition of images. Also, Section 2 gives details of the methodology followed during this study. Experimental results are given in Section 3. At last, Section 4 concludes the paper.

II. MATERIAL AND METHOD

This study consists of three phases. In the first stage, the datasets were collected from four sources. Then we use the image augmentation techniques for improving the training process efficiency. Finally, we apply the pre-trained ResNet50 model of CNN to extract deep features on chest x-ray images.

A. Dataset

CT images of patients labeled Covid19 were used. A data set was created using many sources. The data sets used are as follows; Italian Society of Medical Radiology (SIRM) dataset [11], Corona Virus open-source shared data set [12], data set created by compiling diagnosed images from articles [13], Chest x-ray images data set [14]. An example of Covid-19 positive and negative chest x-ray images from the data set is given in Figure 1.
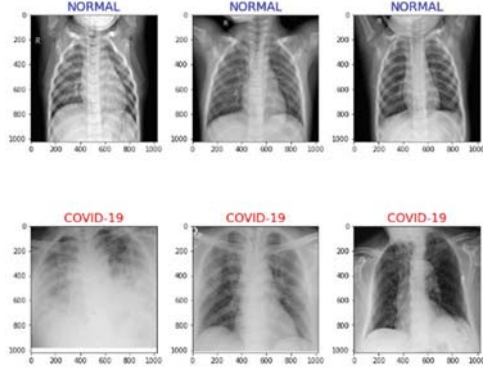

Figure. 1. Covid-19 and Normal (Non-Covid-19) Images

There are two different classes in the data set (Covid-19 and Non-Covid-19). Pixel density distribution to these two different classes is given in Figure 2. If we examine the pixel density distribution graph, we can observe that the distribution of the Covid-19 image class spreads over a wider region.
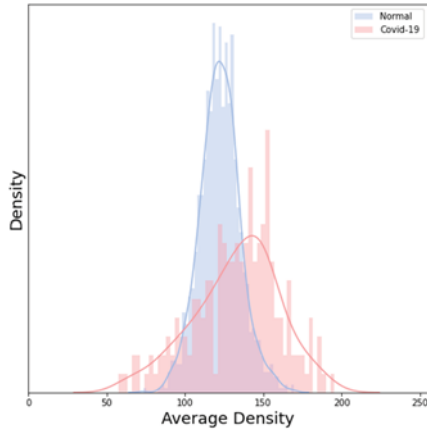

Figure. 2. Pixel Density Distribution of Covid-19 and Normal (Non-Covid-19) Images

B. Image Augmentation Techniques

Image augmentation is a useful technique to duplicate images with some kind of variation so the model can learn from more data. Variations; reflection, scaling, rotation, cutting, and noise are among the commonly used methods [15]. It is based on the expansion of the existing data set by some methods in cases where the number of data is not sufficient during the training phase in the convolutional neural networks. The images are given in Figure 3, after applying image augmentation techniques on our images.
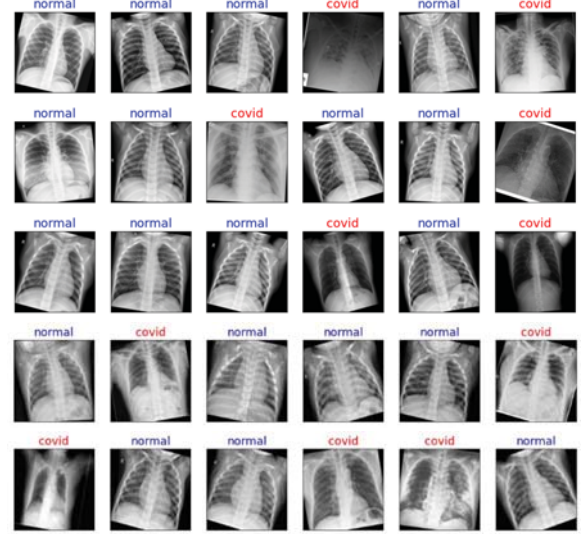

Figure. 3. Covid-19 and Normal (Non-Covid-19) Images After Image Augmentation

C. Convolutional Neural Networks (CNN)

Deep learning is widely used in many fields such as detection of anomalies, object recognition, classification in the biomedical field. One of the most used deep learning architectures is convolutional neural networks. There are convolution, pooling, fully connected layer, and classification layers in convolutional neural networks. Firstly, input data is taken, then it performs the learning process by extracting the features in each layer. It continues with new feature extraction on each layer.

The successful evolution of the convolutional neural network is directly related to both good hardware and a high amount of data. Depending on the hardware, there is a gain in terms of time. It affects the performance result along with the features extracted depending on the amount of data. Transfer learning has been pre-trained on large data sets and can be defined as the use of models by making various adjustments on the current problem by freezing their weights.

ResNet-18 (Residual Network) architecture was developed by the Microsoft research team with the ImageNet dataset to solve the problem of convergence of derivatives to zero which transfer learning. The ResNet-18 model is trained on more than 1 million images, 1000 class data in the ImageNet dataset. ResNet is a model that reduces training time and prevents deterioration with the deepening of the network. Thanks to residual blocks, the input value is added to the result in the layers after it. A general representation regarding the addition of these blocks, input(x), and output are given in Figure 4. [16,17].
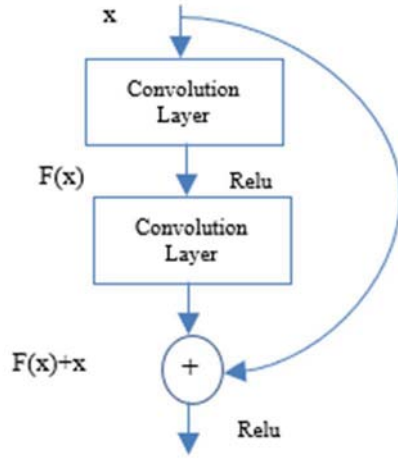
Figure. 4. Residual block [16]

ResNet-50 model, which was pre-trained on the ImageNet dataset, applied to the dataset. The hyperparameters used are as follows; batch size: 32, learning rate: 0.0003, and adam optimizer function. The results were evaluated by applying 5-fold cross-validation on the dataset.

III. RESULTS

As a result of the experiments, we achieved 99.5% classification accuracy. The confusion matrix after classification is given in Figure 5. As the figure shows, only 2 out of 200 Covid-19 images are included in the normal (false negative) class, and only 4 out of 1000 normal (non-covid-19) images are classified as Covid-19 (false positive).

Figure. 5. Confusion Matrix of Covid-19 and Normal (Non-Covid-19) Images

False positive and true positive values are important parameters for performance in classification. Receiver Operating Characteristic (ROC) curve, another criterion, is also used in performance evaluation. The ROC curve is created by plotting false positive values against true positive values. The ROC curve of the classification is given in Figure 6.
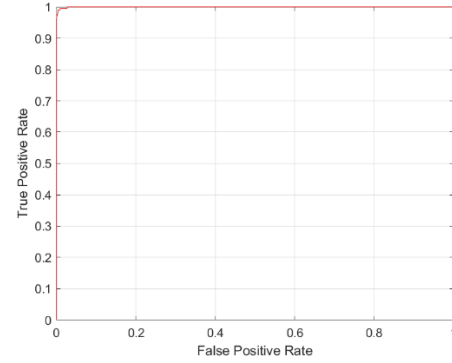

Figure 6. ROC Curve

IV. CONCLUSION AND DISCUSSION

Early diagnosis is essential both for early intervention to the patient and to prevent the risk of transmission. For this purpose, chest x-ray images were used obtained from Covid-19 and non-Covid-19 patients. These images are classified using the transfer learning model ResNet-50. Given that the classification accuracy is calculated with a high accuracy rate of 99.5%, it can help in clinical practice. Also, although the data set is small, the results are encouraging in terms of the use of computer-aided in the field of pathology. It can also be used in situations where the possibilities are insufficient (RT-PCR test, doctor, radiologist).

In future work, more successful deep learning models can be created. Moreover, it can work with larger data sets.

REFERENCES

[1] World Health Organization, and World Health Organization (WHO). "Pneumonia of unknown cause–China', Emergencies preparedness, response." *Disease outbreak news, World Health Organization (WHO) (2020).*

[2] Wu, Fan, et al. "A new coronavirus associated with human respiratory disease in China." *Nature 579.7798 (2020): 265-269.*

[3] Ozturk, Tulin, et al. "Automated detection of COVID-19 cases using deep neural networks with X-ray images." *Computers in Biology and Medicine* (2020): 103792.

[4] Wang, Wenling, et al. "Detection of SARS-CoV-2 in different types of clinical specimens." *Jama* 323.18 (2020): 1843-1844.

[5] Bernheim, A., et al. "CT imaging features of 2019 novel coronavirus (2019-nCoV)." *Radiology* 295.1 (2020): 200463.

[6] Souza, Johnatan Carvalho, et al. "An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks." *Computer methods and programs in biomedicine* 177 (2019): 285-296.

[7] Li, Lin, et al. "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT." *Radiology* (2020).

[8] Abbas, Asmaa, Mohammed M. Abdelsamea, and Mohamed Medhat Gaber. "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network." *arXiv preprint arXiv:2003.13815* (2020).

[9] Wang, Linda, and Alexander Wong. "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images." *arXiv preprint arXiv:2003.09871* (2020).

[10] Apostolopoulos, Ioannis D., and Tzani A. Mpesiana. "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks." *Physical and Engineering Sciences in Medicine* (2020): 1.

[11] Italian Society of Medical Radiology (SIRM) COVID-19 dataset: https://www.sirm.org/category/senza-categoria/covid-19/ (July 2020)

[12] J. C. Monteral. (2020). COVID- Chest x-ray Dataset: https://github.com/ieee8023/covid-chestxray-dataset (July 2020)

[13] Radiopedia: https://radiopaedia.org/search?lang=us&page=4&q=covid+19&scope= all&utf8=%E2%9C%93 (July 2020)

[14] P. Mooney. (2018). Chest x-ray Dataset: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (July 2020)

[15] Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." *arXiv preprint arXiv:1712.04621* (2017).

[16] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[17] Wu, Zifeng, Chunhua Shen, and Anton Van Den Hengel. "Wider or deeper: Revisiting the ResNet model for visual recognition." *Pattern Recognition* 90 (2019): 119-133.