Automatic COVID-19 and Pneumonia Detection from X-ray Images

Serhan Türkan
Graduate School of Sciences and Engineering
Koç University
Istanbul, Turkey
sturkan20@ku.edu.tr

***Abstract*—The COVID-19 pandemic is creating huge outbreaks in over 150 nations, affecting many people's health and lives. One of the most important steps in controlling COVID-19 is early detection and treatment of affected patients. One of the fastest ways to diagnose patients is by using X-ray scans. Early studies found particular anomalies in COVID-19 individuals' chest radiographs. In this study, deep learning algorithms were used to identify people with COVID-19 or other viral pneumonia diseases from chest radiographs. To accomplish this, two Convolutional Neural Network architectures (ResNet18 and ResNeXt101) were used, along with 21 585 chest X-ray images. At the end of the study, ResNeXt101 performed slightly better performance and achieved 94.43% test accuracy. While the achieved performance is somewhat encouraging, additional studies are required on a larger sample of COVID-19 images to provide more accurate estimates.**

***Keywords*—Deep learning, CNN Algorithm, Chest X-ray, Neural Network, COVID-19**

I. INTRODUCTION

Since December 2019, a new coronavirus (SARS-CoV-2) has spread from Wuhan to all of China and then to all countries. According to the report published by the World Health Organization (WHO) on July 20, 2021, coronavirus disease 2019 (COVID-19) has infected roughly 190 million people worldwide and caused more than four million deaths so far [1]. COVID-19 is a virus that affects the lungs and causes tissue damage in people who were affected. Some people may not experience any symptoms in the early stages, whereas the majority of patients experience fever, breathing difficulty, lung infection, and coughing as their primary symptoms [2].

The real-time reverse transcription-polymerase chain reaction (RT-PCR) technique is now the most widely used COVID-19 diagnostic method. The detection procedure, on the other hand, is time-consuming, and the diagnosis result has a significant false-negative rate. Meanwhile, chest imaging examinations such as computed tomography (CT) and X-ray are critical in the early detection of the disease [3].

According to studies, ground-glass opacities (GGO) in a peripheral distribution with bilateral lung involvement are the most prevalent chest X-ray finding in patients. The opacities have a lower lobe inclination, with the right lower lobe being more prevalent than the left lower lobe [4]. Only skilled radiologists can interpret such minor anomalies. Considering the large number of suspects and the limited number of specialist radiologists, the diagnostic time may be longer. Automated techniques to identify such minor anomalies can aid the diagnostic procedure and increase the rate of early detection with high accuracy. Artificial intelligence solutions have the potential to be quite effective in resolving such issues.

Artificial intelligence (AI) approaches were previously used to accurately detect pneumonia from chest X-ray pictures. In the artificial intelligence approach, there are various classification methods. Previously, feature extraction and classification were done manually using classical machine learning methods. Today, we can predict COVID-19 disease directly from raw images without the need for feature extraction with a deep learning method. In recent years, Convolutional Neural Network (CNN), one of the deep learning-based models, have outperformed classical AI approaches in many computer vision processes, including medical image analysis. Moreover, CNNs are multilayer neural networks that are capable of recognizing image patterns without the need for extensive image preprocessing [5].

The goal of this study is to use deep learning methods to classify healthy patients, COVID-19 cases, and viral pneumonia cases from X-ray images. In this study, two different CNN architectures were used to achieve this goal. ResNet18 and ResNeXt101 architectures were trained with the two different dataset and their performances were compared. According to the findings of this study, the architecture with the highest accuracy rate can be used for rapid diagnosis of COVID-19 patients.

This study consists of six sections. Related studies and literature reviews are presented in Section II. Section III describes the methods used to approach the problem. Section IV contains detailed descriptions of the experiments and the parameters used. The results and analysis of this project are presented in Section V. Finally, the findings of this study are reported in Section VI.

II. RELATED WORK

Recently, deep learning approaches have begun to be used instead of classical machine learning techniques. This is because it makes high-accuracy predictions by performing end-to-end processing without the need for manual feature extraction. Therefore, today, the deep learning method is frequently used with the CNN model for image recognition [6].

Several deep learning studies have been conducted to detect COVID-19 from X-ray and CT images for about one-and-a-half-year period since the pandemic began. The primary limitation of these studies was a shortage of the COVID-19 dataset. The X-ray images of COVID-19 cases, whose number did not exceed a hundred at the beginning of the pandemic, reached thousands at the time of this study. These data obtained from patients can be used for higher accuracy training.

Chowdhury et al. used chest X-ray images to create the PDCOVID Net, a novel framework based on parallel-dilated CNN. The authors proposed an approach that utilized a dilated convolution in the parallel stack to capture and stretch essential features, resulting in a detection accuracy of 96.58% [7].

Khan et al. used pre-trained deep learning models such as ResNet18, VGG16, VGG19, and DensNet121 to build a novel architecture for diagnosing X-ray pictures as COVID-19 or normal. VGG16 and VGG19 demonstrated the best accuracies and these models included two phases. Preprocessing and data augmentation are the first phases, followed by transfer learning, which achieves a final accuracy of 99.3% [8].

On X-ray and CT scan image datasets, Maghdid et al. utilized a Convolution Neural Network and transferred learning with Alex Net. CNN had a 94% accuracy rate (88% detail), whereas Alex Net had a 98% accuracy rate (96% detail) [9].

Due to an imbalance in data manipulation and the lack of necessary retrieved features from the images, Ahsan et al. used a fusion of features. HOG and CNN were used to extract these features and classify them in order to improve the detection accuracy of the COVID-19 in their study. With this proposed feature fusion system, they achieved 99.49% accuracy which is the best classification accuracy compared to the other classification techniques like ANN, KNN, and SVM [10].

In this study, a deep learning method was used to detect not only the COVID-19 cases, but also the non-COVID-19 viral pneumonia cases. ResNet18 and ResNeXt101 were used to achieve the best accuracy.

III. APPROACH

A. *Deep Learning*

Deep learning, also known as deep neural network, is an artificial intelligence (AI) function that mimics how the human brain processes data and generates patterns in order to make predictions. Deep neural networks are composed of multiple layers of connected nodes, each layer refining and optimizing the prediction or categorization. This progression of computations through the network is called forward propagation. Visible layers refer to the input and output layers of a deep neural network. The input layer is where the data is fed into the deep learning model for processing, and the output layer is where the final prediction or classification is made [11].

B. *Convolutional Neural Networks (CNN)*

CNNs are one of the most effective algorithms for visual content comprehension and segmentation. They perform well in several computer vision and image processing tasks such as classification, detection, and retrieval of images. The deep CNN's remarkable learning capacity is mainly due to the utilization of several extraction phases that are capable of automating data representations. The availability of extensive data and hardware technology enhancements in CNNs has been speeding up research and new reports have been released on promising deep CNN architectures [12].

C. *Transfer Learning Approach*

Transfer learning (TL) is a machine learning (ML) research subject that focuses on transferring knowledge gained while solving one problem and transferring it to another but similar problem. Pre-trained models are typically trained on large datasets, which are a common benchmark in the computer vision field. The weights computed by the models can be used for various computer vision applications. These models can be used to make predictions on new tasks directly or as part of the training process for a new model. For example, on a smaller dataset, an image classification model trained on ImageNet with millions of tagged photos can be utilized to start task-specific learning for COVID-19 detection. When pre-trained models are used in a new model, the training time and generalization errors are reduced [13].

In this study, since transfer learning produces better results in less time, pre-trained models were used to reduce training time. ResNet18 and ResNeXt101 architectures were used as feature extractors by fine-tuning only the last layer of the Convolution Neural Network

D. *ResNet*

ResNet (Residual Network) was built by He et al., and it was the winner of the ILSVRC 2015. In comparison to earlier networks, their goal was to construct an ultra-deep network that was free of the vanishing gradient issue. Depending on the number of layers (beginning with 34 and going up to 1202), different forms of ResNet were created. ResNet50 was the most frequent form, with 49 convolutional layers and a single FC layer [14].

In comparison to the highway network, ResNet provided shortcut connections inside layers to enable parameter-free and data-independent cross-layer interconnection. When a highway network's gated shortcut is closed, the layers describe non-residual functions. Individuality shortcuts, on the other hand, are never closed, and residual information is always sent through ResNet. Furthermore, because the shortcut connections (residual links) speed up deep network convergence, ResNet has the ability to prevent gradient diminishing difficulties [12]. In this study, 18-layer ResNet18 architecture was used.

E. *ResNeXt*

Xie et al. presented a simple, highly modularized network architecture for image classification. This network is created by repeatedly iterating over a building block that aggregates a collection of transformations with the same topology. Their straightforward design results in a homogeneous, multi-branch architecture with a minimal number of hyper-parameters to configure. This strategy reveals an additional dimension, which we refer to as "cardinality" (the size of the set of transformations), as a critical factor in addition to the depth and width dimensions. They demonstrate empirically on the ImageNet-1K dataset that increasing cardinality can improve classification accuracy even under the restrictive condition of maintaining complexity. Additionally, increasing cardinality is more effective than increasing capacity by going deeper or wider. Their ResNeXt models served as the foundation for our entry into the ILSVRC 2016 classification task, which we won second place. They then investigate ResNeXt on an ImageNet-5K and COCO detection datasets, demonstrating that it outperforms its ResNet counterpart [15]. In this study, 101-layer ResNext101 architecture was used.

IV. EVALUATION/EXPERIMENTS

A. *Dataset*

In this study, two different datasets were used, one for training and validation, the other one for testing, to report cross-dataset performance. The COVID-19 Radiography Database (Version four), the winner of the COVID-19 dataset award by the Kaggle community, is used as training dataset. It contains 3616 COVID-19 positive cases along with 10 192 normal, 6012 lung opacity (Non-COVID lung infection), and 1345 viral pneumonia images [16]. This data set is randomly divided into two parts: 80% as training set and 20% as validation set.

The second dataset, Chest X-Ray (Covid-19 & Pneumonia), which collects images from various publicly available sources, is used as a test dataset. The data set consists of 6432 images, 1583 normal, 4273 pneumonia and 576 covid positive [19].

B. *Preprocessing*

Resizing: The dataset contains X-ray images with different shapes and resolutions. All photos are resized to a fixed size of 224 x 224 pixels to equalize the input dataset and ensure uniformity across all datasets.

Data augmentation: Overfitting, which causes the training accuracy to be higher than the accuracy in the validation/test set, is a common problem in image preprocessing. One way to avoid this problem is to do data augmentation. Data augmentation is a method of increasing the amount of training data by adding modified data from existing data [17]. For this purpose, the images in the dataset were flipped horizontally and rotated, making the dataset larger and more diverse.

C. *Model Hyper-parameters*

Cross-entropy loss, also called logarithmic loss, is used when adjusting model weights during training. Each projected class probability is compared to the actual desired output value of zero or one. Furthermore, a score/loss is calculated that penalizes the probability of deviating from the expected value. The objective is to minimize loss; the less the loss, the more accurate the model. Cross-entropy loss is zero for a perfect model [18].

Both models has been fine-tuned for five epochs. The batch size is set to 8 and 32, and the loss function is optimized using the ADAM optimizer with a learning rate between 0.1 to 0.00001. All applications were made using the PyTorch library.

D. Evaluation Metrics

There are numerous measures for assessing the performance of classification algorithms, including learning models. In this study, the following criteria were used to evaluate the classification results:

- **Accuracy**: The primary criterion for evaluation in the study. It is the ratio of correct predictions to the total number of predictions tested [21].
- **Loss:** The probabilities or uncertainty associated with a prediction based on how far the prediction deviates from the true value [21].
- **Precision**: The ratio of correct positive predictions to the total predicted positives [21].
- **Recall:** The ratio of correct positive predictions to the total positive examples [21].
- **F-Score:** This variable represents the harmonic mean of the recall and precision values [21].
- **Confusion matrix:** The comparison of target estimates with actual values to evaluate classification performance [21].

V. RESULTS AND DISCUSSION

Deep learning methods with CNN architectures ResNet18 and ResNeXt101 are applied to the COVID-19 Radiography Database and Chest X-ray (Covid-19 & Pneumonia) datasets. PyTorch Python libraries were used in the study. Adam was applied as the optimizer algorithm for stochastic gradient descent in two deep learning models. All experiments are performed on the online platform Google Colab, on the GPU, using the Jupyter environment as the backend.

In the first part of the experiment, different hyper-parameters were tested in the models and the training and validation accuracy was observed. The following parameters were tested by returning the model parameters to their initial state every 5 epochs.

Batch size: 8 and 32

Learning Rate: 0.1, 0.01, 0.001, 0.0001 and 0.00001.

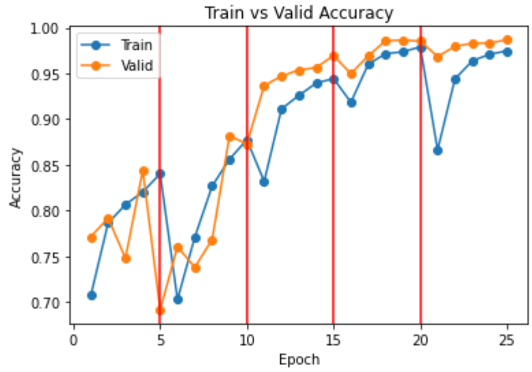

Fig. 1 Accuracy of ResNet18 with batch size 8

Fig. 1 shows the training and validation accuracies according to the learning rate of the ResNet18 architecture trained with eight batch sizes.

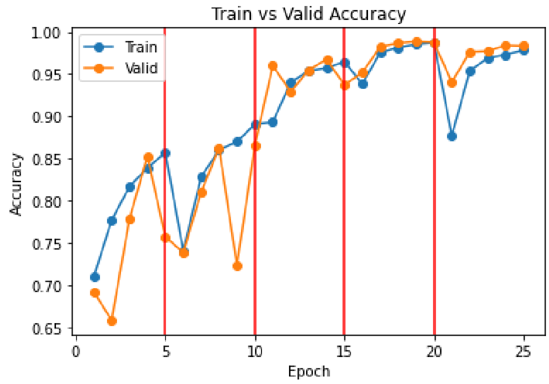

Fig. 2 Accuracy of ResNet18 with batch size 32

Fig. 2 shows the training and validation accuracies according to the learning rate of the ResNet18 architecture trained with 32 batch sizes.

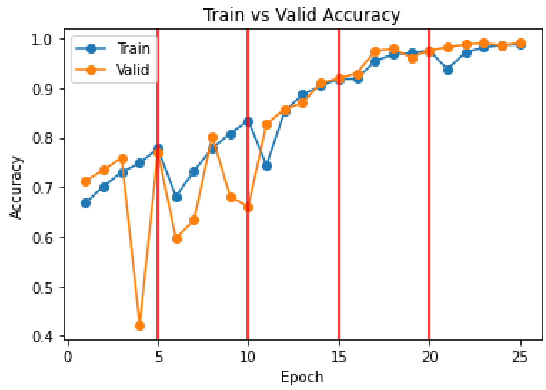

Fig. 3 Accuracy of ResNeXt101 with batch size 8

Fig. 3 shows the training and validation accuracies according to the learning rate of the ResNeXt101 architecture trained with eight batch sizes.

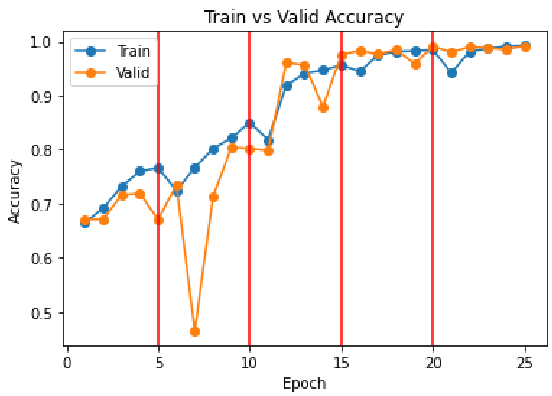

Fig. 4 Accuracy of ResNeXt101 with batch size 32

Fig. 4 shows the training and validation accuracies according to the learning rate of the ResNeXt101 architecture trained with 32 batch sizes.

In the second part of the experiment, both architectures were trained by selecting the hyper-meters that gave the best results in the first experiment. Both models were trained for 20 epochs with a learning rate of 0.00001 and batch size of 32, and the best model parameters were saved according to the validation loss score.

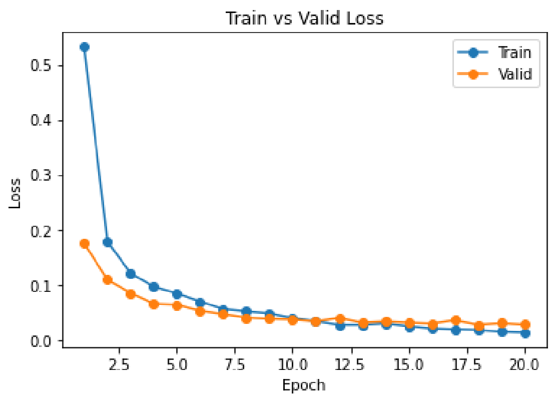

Fig. 5 Training and validation loss of ResNet18

Fig. 5 shows the training and validation loss of ResNet18. Best validation loss score was 0.0268.

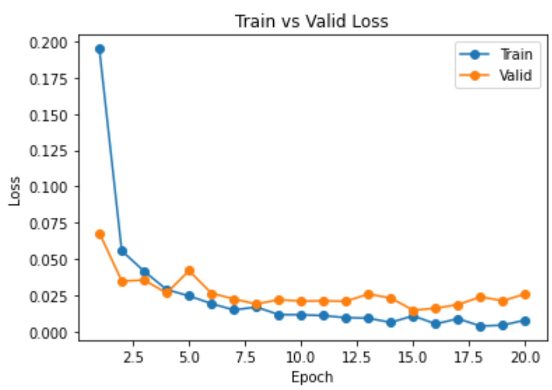

Fig. 6 Training and validation loss of ResNeXt101

Fig. 6 shows the training and validation loss of ResNeXt101. Best validation loss score was 0.0147.

In the final part of the experiment, evaluation was made on the test set with the best models recorded. The accuracy rate was 93.22% for ResNet18 and 94.43% for ResNeXt101. Fig. 7 and 8 show the confusion matrix of the models.

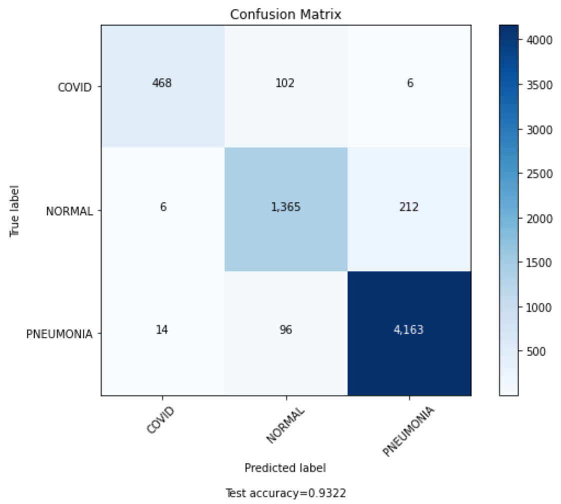

Fig. 7 Confusion matrix of ResNet18

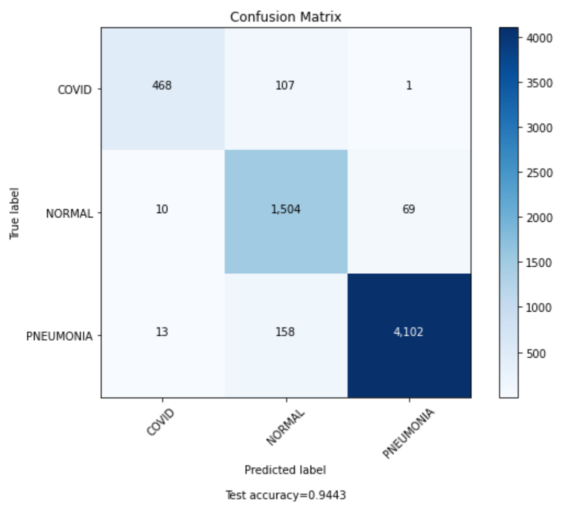

Fig. 8 Confusion matrix of ResNet18

VI. CONCLUSION

The coronavirus pandemic has strained every country's healthcare system to the breaking point since they have had to deal with a significant number of deaths. Early detection of COVID-19 in a more rapid, simple, and cost-effective manner can help save lives and alleviate the burden on healthcare professionals. Artificial intelligence, specifically CNN, can play a significant role in detecting COVID-19 by applying image processing techniques to X-ray images.

In this study, a total of 21 165 chest X-ray images, 4192 of which are COVID-19 positive, were analyzed using deep learning methods for the diagnosis of COVID-19. In the experiment using the same hyper-parameters, ResNeXt101 achieved 94.43% accuracy and ResNet18 achieved 93.22% accuracy. Despite being a much smaller architecture, ResNet18 achieved a result close to the much larger layered ResNeXt101.

In the future studies, the model can be evaluated for CT imaging. By building various deep learning models, one can improve their success rate and performance. When the number of X-ray images and other medical images continues to grow, deep learning technologies will be applied in the health industry on a daily basis. Patients who test positive for COVID-19 can be admitted immediately without waiting for lab test results, hence increasing the survival rate.

References

[1] [Online]. Available: https://www.who.int/publications/m/item/weekly-operational-update-on-covid-19---20-july-2021

[2] Sheikhi, Kowsar, Hamidreza Shirzadfar, and Milad Sheikhi. "A review on novel coronavirus (Covid-19): symptoms, transmission and diagnosis tests." Research in Infectious Diseases and Tropical Medicine 2.1 (2020): 1-8.

[3] Zu, Zi Yue, et al. "Coronavirus disease 2019 (COVID-19): a perspective from China." Radiology 296.2 (2020): E15-E25.

[4] [Online]. Available: https://bmcpulmmed.biomedcentral.com/articles/10.1186/s12890-020-01286-5

[5] Qin, Chunli, et al. "Computer-aided detection in chest radiography based on artificial intelligence: a survey." Biomedical engineering online 17.1 (2018): 1-23.

[6] Minaee, Shervin, et al. "Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning." Medical image analysis 65 (2020): 101794.

[7] Chowdhury, Nihad K., Md Muhtadir Rahman, and Muhammad Ashad Kabir. "PDCOVIDNet: a parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images." Health information science and systems 8.1 (2020): 1-14.

[8] Khan, Irfan Ullah, and Nida Aslam. "A deep-learning-based framework for automated diagnosis of COVID-19 using X-ray images." Information 11.9 (2020): 419.

[9] Maghdid, Halgurd S., et al. "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms." Multimodal Image Exploitation and Learning 2021. Vol. 11734. International Society for Optics and Photonics, 2021.

[10] Ahsan, Mominul, et al. "COVID-19 detection from chest X-ray images using feature fusion and deep learning." Sensors 21.4 (2021): 1480.

[11] [Online]. Available: https://www.ibm.com/cloud/learn/deep-learning

[12] Khan, A., et al. "A survey of the recent architectures of deep convolutional neural networks. arXiv." Preprint 10 (2019).

[13] [Online]. Available: https://neptune.ai/blog/transfer-learning-guide-examples-for-images-and-text-in-keras

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

[15] Xie, Saining, et al. "Aggregated residual transformations for deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[16] [Online]. Available: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

[17] Shorten, C., and T. M. Khoshgoftaar. "A survey on image data augmentation for deep learning. J. Big Data 6 (1), 1–48 (2019)."

[18] [Online]. Available: https://machinelearningmastery.com/cross-entropy-for-machine-learning/

[19] [Online]. Available: https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia

[20] [Online]. Available: https://docs.paperspace.com/machine-learning/wiki/accuracy-and-loss

[21] Sevi, Mehmet, and İlhan Aydin. "COVID-19 Detection Using Deep Learning Methods." 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI). IEEE, 2020.