A Novel Approach to Detect Coronavirus Based on Lung x-ray Image and Convolutional Neural Networks

Huda Mubarak Ismail[*1], Pedram Salehpour[2], Seyed Hadi Aghdasi Alamdari[3]

[1]Ministry of Defense, Baghdad, Iraq
[2,3]Department of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
Emails: hudamod2017@gmail.com ; Psalehpoor@tabrizu.ac.ir; aghdasi@tabrizu.ac.it

Abstract
The coronavirus has become a global crisis in recent months. The virus has disrupted or shut down many social, economic, sports, scientific, etc. activities. In addition to the medical importance of this Disease, its rapid and accurate diagnosis is an important need. In this study, we proposed a novel method to detect coronavirus using machine learning and classification algorithms based on lung images. In general, the method consists of two steps. At the first step, a convolutional neural network is trained with a data set of lung images that determine whether a viral infection exists or not. In the second step, another network is used to detect if the existence of viral infection is considered coronavirus or not. Experimental tests have been conducted that show the correct diagnosis can be made with 95.8%.

Keywords: Corona diagnosis; deep learning; convolutional neural networks; lung images.

1. Introduction

In early 2020, a mysterious virus called Corona underwent many changes in the world. It is safe to say that the Coronavirus has become a global crisis in recent months. The virus has disrupted or shut down many social, economic, sports, scientific, etc. activities. The effects of the virus on the health and economy of the world are such that some thinkers have divided the world into two periods before the corona and believe that the world will change a lot after the corona challenge.
In this study, our main goal is to build a coronavirus detection system that has three main features:
• Be cheap and affordable: This will be an important feature for this system. Considering this feature, everyone can use it.
• Be fast: The corona detection system must announce the result quickly.
• Accurate: The corona detection system must have an acceptable percentage of accuracy.

2. A review of related works

In this part, some related works will be reviewed, [1] propose to use two known deep learning networks, SegNet and U-NET, for image tissue classification. The results in [1] show the superior ability of SegNet in classifying infected/non-infected tissues compared to the other methods (with 0.95 mean accuracy), while the U-NET shows better results as a multi-class segmentor (with 0.91 mean accuracy). Paper [2] reports a convolutional neural network (CNN) based multi-image augmentation technique for detecting COVID-19 in chest X-Ray and chest CT scan images of coronavirus suspected individuals. See Figure 1.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

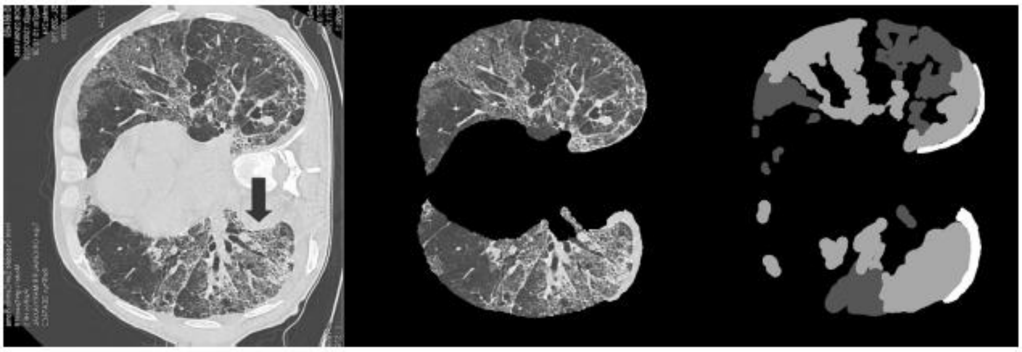


Figure 1: Dataset sample used in [1]. CT scan (left), masked lungs (middle), and labeled classes (right), where black is class C0, dark gray is C1, light gray is C2, and white is C3 [1]

Narin and etl. [3] have implemented three different binary classifications with four classes (COVID-19, normal (healthy), viral pneumonia, and bacterial pneumonia) by using 5-fold cross-validation. Considering the performance results obtained, it has been seen that the pre-trained ResNet50 model provides the highest classification performance (96.1% accuracy for Dataset-1, 99.5% accuracy for Dataset- 2, and 99.7% accuracy for Dataset-3) among other four used models.

In this study, a novel method is proposed that unlike other methods it uses two networks to classify the patients. The first network determines whether the patient has a viral infection at all or not. In case the first network does not detect any viral infection, the patient will be assumed uninfected of Covid-19. Otherwise, the second network will determine if the viral infection is Covid-19 or other viral infections. The proposed method shows that the accuracy of the proposed method is reliable and it can be considered as a new approach in addition to other available methods to detect Covid-19.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

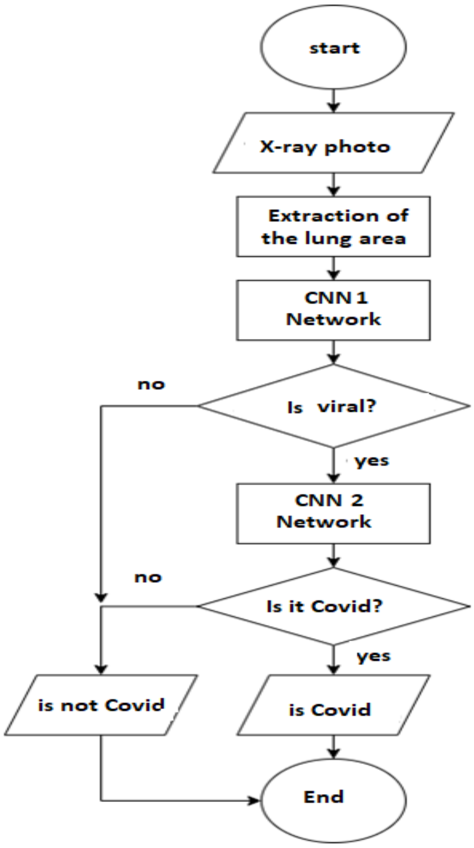


Figure 2: Flowchart of the proposed method

3. The proposed method

Major headings should be typeset in boldface with the first letter of important words capitalized.
As mentioned, identifying a patient with Covid from lung images is one of the safest methods of diagnosis. The basic block diagram of the proposed method is shown in Figure 3.

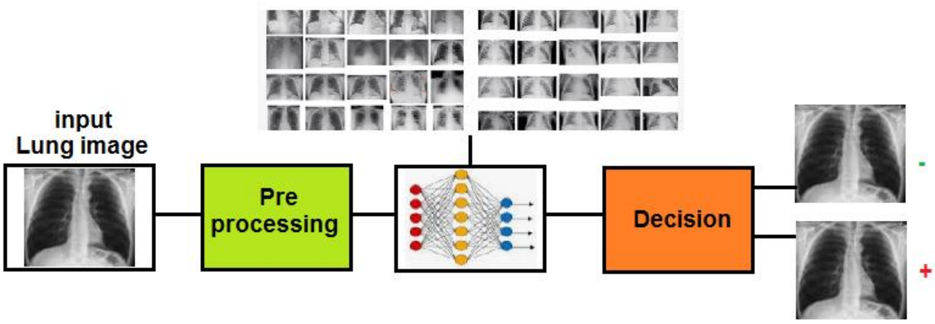


Figure 3: Outline of the proposed system

To achieve an automatic diagnosis of the Disease by the machine, we divide the work into three general parts:

- Database preparation
- Pre-processing and diagnosis of the pulmonary region
- Identify the infected person

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

3.1. *Database preparation*

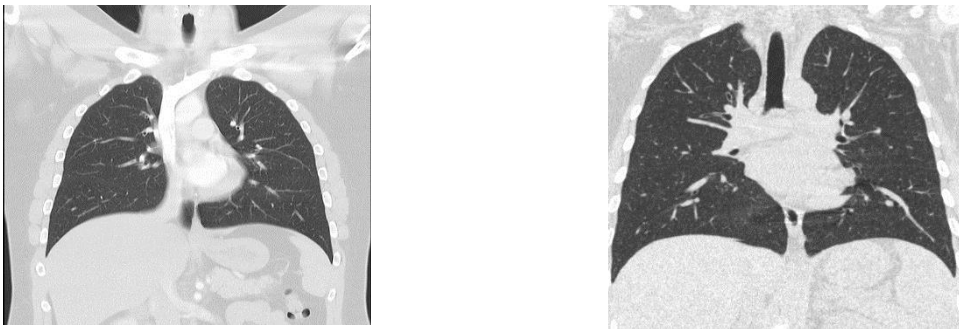
The most important part of machine learning is access to a rich database with a wide variety of images of the subject under study. This database should contain 3 groups of lung images:

a. Images of a healthy person's lungs
b. Images of a patient with lung Disease
c. Lung images of a patient with a lung Disease other than Covid

Like other CT scans, the lungs use special CT scan equipment to diagnose possible problems and abnormalities. CT scans of the lungs are usually given when the patient has symptoms such as unexplained cough, chest pain, shortness of breath, fever, and so on.

3.2. *Pulmonary area detection*

One of the main parts of image processing is the removal of background and unnecessary items from the images under review. In the introduced work, the lung area should be extracted well from the photo. This section shows its importance in diagnosing the lung area for infected people. In the lung images of healthy people, the contrasting color of the lung area is clearly visible in the images. However, in people with lung diseases, due to the presence of damaged tissues, the color of the lungs becomes the same as other areas of the body and it becomes impossible to diagnose the lungs with old methods. In this work, it is suggested that CNN deep neural network be used to diagnose the lung. A CT scan of a healthy lung of a 25-year-old man and a CT scan of the lung of a 30-year-old woman are shown in the figures below.



(a) (b)

Figure 4: (a) Healthy lung image of a 25-year-old man (b) CT scan of the lungs of a 30-year-old woman

In these CT scans, contrast is used for better display. In these pictures, it is clear that the lung tissue is normal. The axial image of the lung shows that the lung is completely clean and there are no complications or abnormalities.
A normal, healthy corona lung should look black on a CT scan. It is usually normal to have small lumps of tissue or lung nodules that appear as tiny white spots.
But a coronary lung CT scan has white spots that radiologists refer to as "opaque glass opacity". Figure 5 shows an image of the corona in the lung health test by CT scan of the coronary lung. This sign indicates a severe COVID-19 infection caused by fluid buildup in the lungs.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

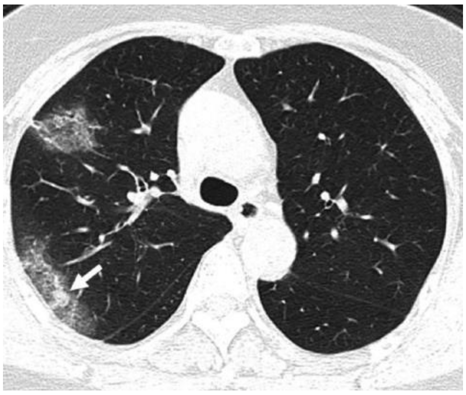


Figure 5: Lung image of a person suspected of having a coronary artery

3.3. *Identifying the infected person*

With the help of the previous steps, the extracted area is used for diagnosis. After studying, we concluded that the tissues of Covid Disease appear reticular in the lungs. With the help of this knowledge, it can be concluded that a violet conversion is a good option for extracting basic features. Violet conversion is one of the most widely used mathematical transformations in the field of processing, especially signal and image processing. Due to the nature of the multisolution analysis, this conversion has found its place in many processing applications and sometimes occurs as the most powerful tool. This conversion is highly efficient in pattern recognition in geometric images.

These features, along with the base images, will be given to the new CNN network to be used to diagnose Covid Disease after training with database images. CNN has become popular because of its better performance in image classification. Convection layers in the grid along with filters help to extract spatial and temporal features in the image. These layers have a weight distribution technique that helps you reduce the computational load [4, 5].

Architecturally, CNNs are simply artificial feedback neural networks (ANNs) with two limitations: the neurons in a filter are attached to only part of the image to maintain spatial structure, and their weight is shared to reduce the number of model parameters. Becomes. A CNN is made up of three building blocks: (i) a convolution layer to learn features, (ii) a maximum pool layer (sampling) to lower the image sample and reduce the dimensions, thereby reducing the computational load, and (iii) a fully interconnected layer to equip the network with classification capabilities [6]. The CNN architecture is shown in Figure 6.

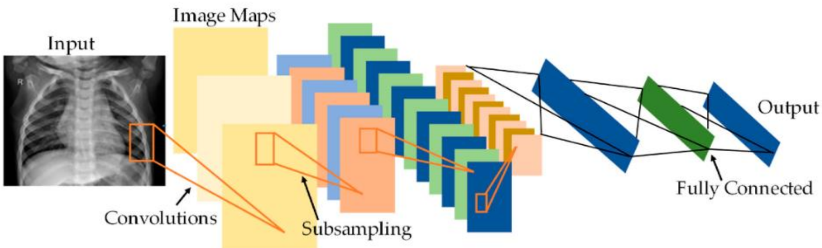


Figure 6: CNN architecture i

Convolution is the main idea behind CNN, and how it is used is a major determinant of network performance. Using networks larger than 3. 3 does not have much of an advantage on CNN, so you should avoid using 5 × 5 or 7 × 7 networks on your network. Consecutively accumulating 3 × 3 convulsions, along with VGGNet and ResNet, have been shown to achieve the same "receptive field" that larger convulsions provide. Computing is also more efficient.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

4. Results

The comparative performance of training and test accuracy for different CNNs for classification schemes is shown in Figure 7. It should be noted that DenseNet201 offers the highest accuracy for training and testing. This accuracy was found to be 93.3%. Figure 8 shows the sub-curve area (AUC) / receiver characteristics curve (ROC) (known as AUROC (area below the receiver performance characteristic)) for different classification schemes, which is one of the most important evaluation criteria to consider. Figure 7 shows that DenseNet201 performs better than other algorithms.

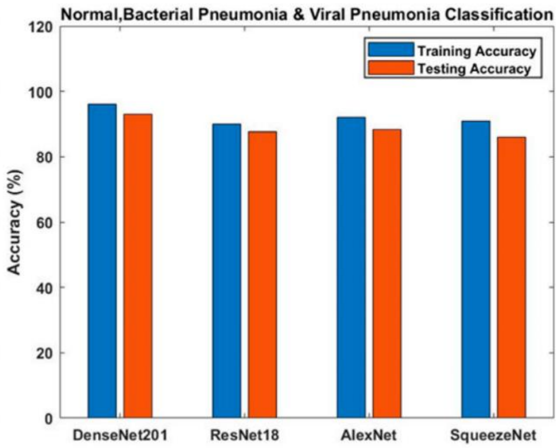


Figure 7: Comparison of training and experiment accuracy for classification using different models

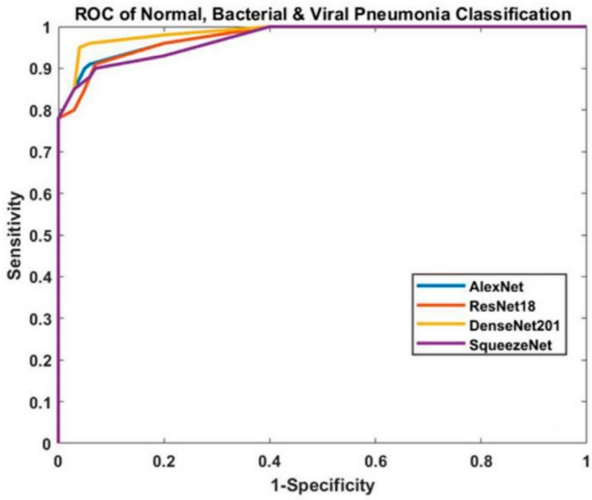


Figure 8: Comparison of receiver performance characteristics (ROC) curves for classification using CNN-based models

The perturbation matrix for the performance of the superior algorithm, which is the same as the pre-trained Densenet201 model, is shown in table 1.

Table 1: Turbulence matrix for classification using DenseNet201.

	Bacterial Disease	**Healthy (natural)**	**Viral Disease**	**Accuracy**	**Error**
Bacterial Disease	168	7	22	85.3 %	**14.7 %**

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

Healthy (natural)	0	195	4	98 %	**2 %**
viral Disease	2	5	194	96.5 %	**3.5 %**

Table 2 summarizes the performance matrices for the various CNN algorithms tested. DenseNet201 performs better than other models in terms of various indicators.

Table 2: Different performance metrics for different deep learning networks.

Models	Accuracy	Sensitivity	Specificity	Precision (PPV)	Area under Curve (AUC)	F1 Scores
AlexNet	0.885	0.911	0.886	0.941	0.883	0.884
ResNet18	0.909	0.91	0.875	0.94	0.88	0.877
DenseNet201	0.933	0.932	0.967	0.937	0.95	0.935
SqueezeNet	0.865	0.895	0.87	0.93	0.859	0.861

It is obvious that DesneNet201 had the highest accuracy. As a result, this model was applied to the required network in both stages.
It should be noted that the accuracy stated in Table 2 is the average accuracy and as can be seen in Figure 7, the accuracy of diagnosing viral Diseases in the first stage of the algorithm is 96.5%. After using the DesneNet201 model for both the existing networks in the algorithm and the training of the second network, table 3 shows the final accuracy of the algorithm's performance in diagnosing Covid Disease.

Table 3: Final accuracy of algorithm performance

	Covid Disease	Other diagnoses	Accuracy	Error
Covid Disease	95	4	95.9 %	**4.1 %**
Other diagnoses	8	182	95.7 %	**4.3 %**

As a result, the final accuracy of the algorithm will be 95.8%. Compared to previous work, the proposed method has been faster and more accurate. It should be noted that the average training time in this method is 14 hours, and the average test of each entry is 7 seconds. Our study is more accurate than methods such as [7] and more rapid and less complex than methods such as [8-14].

5. Conclusions

We believe that this computer-aided diagnostic tool can significantly assist the radiologist in diagnosing lung disease by type. This rapid classification opens up other avenues for using this CAD tool, especially in screening Covid patients at the airport. In future research, better results can be achieved by combining some other neural networks. Data preprocessing can also be done with some other algorithms. Another issue is the quality of lung images. Imaging speed, etc. can all be topics for future research.

References

1. A. Saood and I. Hatem, “COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet,” BMC Med. Imaging, vol. 21, no. 1, 2021.
2. K. Purohit, A. Kesarwani, D. R. Kisku, and M. Dalui, “COVID-19 detection on chest X-Ray and CT Scan images using multi-image augmented deep learning model,” bioRxiv, 2020.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022

3. A. Narin, C. Kaya, and Z. Pamuk, “Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks,” Pattern Anal. Appl., vol. 24, no. 3, 2021.
4. S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017, 2018, vol. 2018-January, pp. 1–6.
5. M. Goyal, R. Goyal, and B. Lall, “Learning Activation Functions: A new paradigm of understanding Neural Networks,” Arxiv, 2019.
6. C. Bailer, T. Habtegerbrial, K. Varanasi, and D. Stricker, “Fast Feature Extraction with CNNs with Pooling Layers,” Arxiv, 2019.
7. V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, “Diagnosis of COVID-19 using CT scan images and deep learning techniques,” Emerg. Radiol., vol. 28, no. 3, 2021.
8. S. H. Abbood, H. N. Abdull Hamed, M. S. Mohd Rahim, A. H. M. Alaidi, and H. T. Salim ALRikabi, "DR-LL Gan: Diabetic Retinopathy Lesions Synthesis using Generative Adversarial Network," *International Journal of Online and Biomedical Engineering(iJOE),* vol. 18, no. 3, 2022.
9. A. S. Abdalrada, O. H. Yahya, A. H. M. Alaidi, N. A. Hussein, H. T. Alrikabi, and T. Al-Quraishi, "A predictive model for liver disease progression based on logistic regression algorithm," *Periodicals of Engineering and Natural Sciences,* Article vol. 7, no. 3, pp. 1255-1264, 2019.
10. I. A. Aljazaery, H. T. Salim ALRikabi, and A. H. M. Alaidi, "Encryption of Color Image Based on DNA Strand and Exponential Factor," *International Journal of Online Biomedical Engineering,* vol. 18, no. 3, pp. 101-113, 2022.
11. M. K. Abdul-Hussein and H. Alrikabi, "Evaluation of the Interference’s Impact of Cooperative Surveillance Systems Signals Processing for Healthcare," *International Journal of Online and Biomedical Engineering(iJOE),* vol. 18, no. 3, pp. 43-59, 2022.
12. H. T. ALRikabi and H. T. Hazim, "Enhanced Data Security of Communication System Using Combined Encryption and Steganography," *International Journal of Interactive Mobile Technologies,* vol. 15, no. 16, 2021.
13. M. Al-dabag, H. S. ALRikabi, and R. Al-Nima, "Anticipating Atrial Fibrillation Signal Using Efficient Algorithm," *International Journal of Online and Biomedical Engineering (iJOE),* vol. 17, no. 2, pp. 106-120, 2021
14. M. Canayaz, “MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images,” Biomed. Signal Process. Control, vol. 64, 2021.

DOI: https://doi.org/10.54216/IJAIET.010102
Received: January 11, 2022 Accepted: May 22, 2022