Review Article

Application of machine learning and medical imaging in the detection of COVID-19 patients: A review article

Sepideh Yadollahi[1], Setareh Yadollahi[2], Elmira Zanjani[3], Fatemeh Khaleghi[4]

[1]M.D., Postdoctoral Research Scholar, Mayo Clinic, Rochester, Minnesota, U.S, [2]Undergraduate Medical Student, Medipol University, Istanbul, Turkey, [3]Department of Medical Physics and Biomedical Engineering, School of Allied Medical Sciences, Tehran University of Medical Sciences (TUMS), Tehran, [4]Assistant Professor, Department of Radiology, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

Abstract

In the present study, a particular technique of artificial intelligence (AI) is applied for diagnosis and classifying medical images of patients with coronavirus disease (COVID-19). Chest radiography and laboratory-based tests are two of the most important diagnostic approaches for the detection of people with the coronavirus. Recently, a lot of studies have been carried out on using AI techniques for achieving appropriate diagnosis of COVID-19 patients using computed tomography (CT) of the chest. The present study is reviewing all available literature that have investigated the role of chest CT toward AI in the detection of COVID-19. As a novel field of computer science, AI focuses on teaching computers to be capable of learning complex tasks and decide about their solution methods. In this study, we used Matlab, Payton, and Fortran software as well as other software which are suitable for this research. In this regard, the present review study is aimed to collect the information from all the studies conducted on the role of AI as a decisive and comprehensive technology for the detection of coronavirus in patients to have a more accurate diagnosis and investigate its epidemiology.

Keywords: Artificial intelligence, COVID-19, machine learning, medical image

Introduction

In December 2019, the pandemic of coronavirus disease (COVID-19) was identified in the city of Wuhan, China, which has an intense public health problem worldwide.[1,2] The cause of this review study (COVID-19) is to investigate severe acute respiratory syndrome (SARS) and its medical presentations. The Hubei region in China was known for two different coronaviruses, which are Middle East respiratory syndrome and SARS.[2] Coronaviruses spread from cold to severe diseases that can affect both humankind and animals. The World Health Organization (WHO) has shown that the total number of cases of coronavirus (COVID-19) infected is 2,995,758, and 204,987 deaths are reported from the coronavirus (COVID-19) pandemic incidence till the 29 April 2020 from all over the world (WHO, 2020). The spread of this disease transmission from human to human is completely fast and going up exponentially. The highest cases reported from outside China are in the United States, Spain, Italy, United Kingdom, Germany, France, Turkey, Russia, Iran, Brazil, and Canada (WHO).

The association of COVID-19 and weakness has been reported by many studies in various countries around the world. Unfortunately, it has also been reported that healthcare systems are unable to manage COVID-19 patients appropriately that caused anxiety among them. It should be noted that the lack of accuracy in clinical methods of detection of COVID-19

Address for correspondence: Dr. Fatemeh Khaleghi, Department of Radiology, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.
E-mail: fateme.khaleghi2016@gmail.com

Received: 25-05-2021 **Revised:** 17-08-2021
Accepted: 13-12-2021 **Published:** 30-06-2022

Access this article online
Quick Response Code:
Website: www.jfmpc.com
DOI: 10.4103/jfmpc.jfmpc_1715_21

How to cite this article: Yadollahi S, Yadollahi S, Zanjani E, Khaleghi F. Application of machine learning and medical imaging in the detection of COVID-19 patients: A review article. J Family Med Prim Care 2022;11:2277-83.

has been rapidly spread worldwide.[3] In this regard, two of the newest clinical techniques for the detection of COVID-19 are introduced to be chest computed tomography (CT) and chest radiograph (X-ray) scans.[3,4] These novel techniques are capable of revealing anomalies in the chest that represent different diseases such as COVID-19. X-ray and CT scan imaging techniques could be applied as primary techniques for assessing COVID-19 severity, forecasting the progression of COVID-19, and monitoring the emergency case of infected patients.[5] In this regard, the application of AI has grown up in various medical detection fields.[6] Nowadays, the application of AI has been increased widely for achieving more accurate and satisfactory results and diminishing the burden on the healthcare system.[7] Through these techniques, the decision time related to the detection process will significantly. The development of AI methods has a significant role in achieving a more accurate diagnosis of the risks of epidemic diseases that provide more precise prediction results, and also the possibility of prevention and detection of any possible health risks in the future.[8] An important technique for early detection and assigning priority to those patients who are suspected to be affected with SARS-CoV-2 infection is chest CT. For achieving a better understanding of the disease pathogenesis for controlling the incidence of infection, a comprehensive collaboration between physicians, virologists, phylogeneticists, epidemiologists, policymakers, and public health official workers should be created. Common imaging patterns could be observed in early investigations through chest CT.[8,9] For instance, at the beginning of the COVID-19 pandemic, a prospective analysis in Wuhan revealed that in 98% of infected patients bilateral lung opacities were observed. Moreover, this technique described subsegmental and lobular regions of lung consolidation as the key findings of imaging.[10,11]

Literature review

Coronavirus originates in China and growth is fast among other countries. The COVID-19 can stay on different surfaces for a few days such as steel, plastic, copper. The newest way is used to diagnose and classify the COVID-19 using support vector regression method to identify five different tasks related to coronavirus. The tasks are to find the spreading of COVID-19 across areas, the growth measure of COVID-19 infected patients and their mitigation, how the epidemic's end, analyzing the transmission rate of the virus, and the correlation of COVID-19 with weather conditions. The finding is compared with other regression models (linear regression, polynomial regression), which provides promising results both in accuracy and efficiency.[1]

Based on WHO, the COVID-19 is a pandemic disease. The X-ray and CT images illustrate detection and infected area localization of COVID-19 from two different modalities of medical imaging. The IRRCNN and NABLA-3 network models have been used for the classification and segmentation process. The finding of the proposed algorithm indicates that 84.6 and 98.78% results have been provided from the X-ray and CT images, respectively.[12] The only specific method to diagnose the virus is by taking a trial. A deep learning way (ResNet-18, ResNet-50) has been used to classify this disease (COVID-19) from CXR images. The COVID-19 CXR data sets (CORDA, RSNA, COVID-chest X-ray, chest X-ray) have been collected from the emergency hospital in Northern Italy. The analysis of outcomes from the proposed deep learning algorithm indicates that the training with a COVID-net architecture and COVID-chest X-ray data set is 85% accuracy and the ResNet-18 architecture with a COVID-chest X-ray data set is 100% accuracy.[13]

The most suitable method of controlling over the coronavirus is quarantine and appropriate treatment. Machine learning algorithm and artificial intelligence (AI) techniques are both used for achieving an appropriate detection of COVID-19 from chest X-ray images. The images were classified using CNN with the SoftMax classifier, SVM, and random forest. Classification of COVID-19 could be done using CNN based on chest X-ray images that mainly uses machine learning techniques for two scenarios; the image is classified and extract the graphical feature for a hybrid system. The extracted feature has been used further for training and testing parameters. Based on the proposed algorithm, the accuracy of CNN is more than 95% that is a prove of the claim that this method is better than others.[14]

A deep-running technique has been introduced for identifying the COVID-19 patients, which is mainly based on X-ray images. X-ray images will be classified from deep features by the supported vector machine (SVM). The suggested classification model of Resnet50 plus SVM is proved to have an accuracy of 95%. All the used data in the present study have been collected from the online repository of GitHub and the Kaggle site.

Detecting the coronavirus (COVID-19) based on radiology and radiography scans are of the most appropriate and efficient techniques for the management of patients with COVID-19. Therefore, a deep learning framework from chest X-ray images has been introduced for the detection of COVID-19 patients. The four pretrained convolutional models, including ResNet18, ResNet50, Squeeze Net, and DenseNet-121, have been used to identify the coronavirus (COVID-19) disease. The two data sets are considered and combined, which have been around 5k images called COVID-X-ray-5k. The result from the all pretrained model shows that sensitivity 97% and specificity 90% are achieved.[15]

- Diagnosis of COVID-19 in suspected patients was carried out by conducting a reverse transcriptase-polymerase chain reaction (RT-PCR) trial on SARS-CoV-2 virus-specific cells. Anyway, the completion of this test may take up to 2 days, and then serial testing may be required to eliminate any possible false-negative outcomes. However, currently, the lack of appropriate access RT—PCR test kits underscore an urgent need for using an optional method to achieve rapid and accurate diagnosis of COVID-19 patients. CT scan is a significant element for the assessment of patients who may suffer from SARS-CoV-2 infection. However, using CT alone may have limited negative predictive value in eliminating

SARS-CoV-2 infection because the radiological results of infected patients at early stages of the disease may be normal. In the present study, we use AI algorithms for integrating chest CT findings with clinical symptoms, laboratory testing, and the history of exposure to infection for achieving a rapid diagnosis of patients who are suspected of COVID-19. In 46.3% of patients tested by real-time RT-PCR examination and next-generation sequencing RT-PCR, the result was positive for SARS-CoV-2. In another study, among 279 tested patients, the results for the AI system achieved an area under the curve of 0.92 with equal sensitivity in comparison with the senior thoracic radiologist. Moreover, the applied AI system improved the detection rate of patients whose RT-PCR test was positive for COVID-19 while their CT scans were normal. For example, 68% of patients who were detected with a negative COVID-19 by radiology scans were correctly identified as positive cases through this method. The availability of clinical history of patients and their CT scans at the same time with using the AI system swift the diagnosis process of COVID-19 patients.[16]

Various radiological studies on COVID-19 characteristics have been published since the American College of Rheumatology recommendations were published. Achieved data showed a considerable improvement in the geographical coverage of this method at the same time with the increment of the accuracy of interpretation. Additionally, novel models of AI have been formed and developed to help radiologists in improving their accuracy in the diagnosis of COVID-19 using radiological evidence. Another study demonstrated that using AI for clarifying radiological data provides the possibility of modifying and improving the diagnosis of COVID-19 in comparison with radiologists draw conclusions without AI-based help.

Methodology

The methodology of this study focuses on the detection and diagnosis of Covid 19 by machine learning and medical images using AI. This review study is aimed to represent a novel method of rotational scaling and a way of translation invariant character extraction for achieving a more precise diagnosis of Covid 19. The main steps of detection of COVID-19 are composed of the following steps:

First of all image acquisition process was conducted, after that, the images were processed through the image preprocessing technique. At the next step, images went through the Binarization process, and their data transformed into the vectors of binary numbers. Then, the images went through the process of partitioning a digital image into multiple segments that is called the segmentation stage. Following were the steps used: thresholding, feature extraction, and neural network detection.

For this purpose and to obtain precisely detailed scans of regions inside the body, a particular type of digital X-ray machine is used. A total of 200 lung CT scans were collected from the Internet and hospital documents of patients with COVID-19 and those with the normal images of lungs. All images used in this system were in a.jpeg file format.

The thresholding technique was used to conduct the image segmentation process which is particularly based on threshold value for converting a gray-scale image to a binary image. Choosing the appropriate threshold values is the key idea of this technique. Nowadays, novel methods have been developed for computing appropriate threshold values. Our proposed system is composed of three types of threshold values known as Thresh 1, Thresh 2, and Thresh 3. In binary CT image, when the introduced Thresh 1 is lower than the percentage of white pixels, the lung is fully affected. Its while, in the segmented binary picture when the percentage of white pixels is greater in comparison with Thresh 2 and Thresh 3, both left and right lungs are affected.

Image acquisition

Images with precise details of areas inside the body were obtained using a special type of digital X-ray machine as a powerful computerized tomography (CT) scanner machine. A total of 300 lung CT scans were collected from the Internet and hospital documents of patients with cancer and those with the normal images of lungs. All used lung CT images in this system were in a.jpeg file format.

Image preprocessing

After completing the process of image acquisition, image preprocessing stars. All the steps of image preprocessing are presented in its block diagram as shown in [Figure 1 and 2].

Segmentation image

In the interdisciplinary scientific field, the computer vision system is a technique in which digital images are partitioned into various segments. The main objective of segmentation process is to decrease the delegation of scans to make their analysis easy and meaningful. The process of locating objects and boundaries in

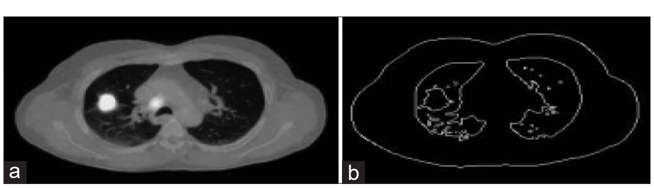

Figure 1: (a) Original grayscale picture and (b) just edge

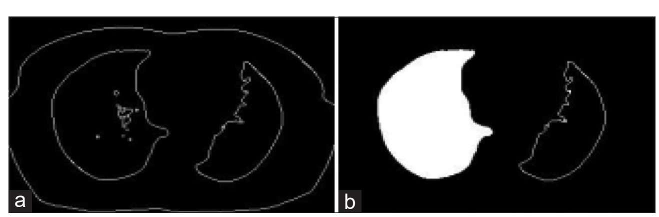

Figure 2: (a) Dilated scan and (b) filled scan

images is carried out by image segmentation. Image segmentation could be carried out more precisely by assigning a label to every pixel in a picture so that the pixel with its symbol shares certain characteristics. Defined segmentation processes consist of different processes as mentioned in the following. As could be seen from [Figure 2], initially, the original gray-level image was converted into an edge just image. After that, the image line image was changed into a dilated image and then a filled image was created. Finally, both left and right lungs were segmented from the image marked as shown in [Figure 3].

Results

Characteristics of study population

Based on the data from Table 1 all the data of population characteristics are presented that could be used as the training and independent testing data sets. Based on the results of the present study, it could be seen that the rate of male patients in groups of COVID-19, community-acquired pneumonia (CAP), nonpneumonia patients is higher than females. In CAP and COVID-19 groups, patients were older compared to the patients in the nonpneumonia group for both data sets of training and independent testing.

Model performance

After training the deep learning pattern, the time of processing for new testing examinations will be more swift. For any CT scan examination, the average processing time was 4.51 s with a workstation. All the tests were carried out with a computer equipped with GPU NVIDIA Quadro M4000, 16GB Ram, and Intel Xeon Processor E5-1620 v4 @3.5GHz. [Figure 3] illustrates the receiver operating characteristic (ROC) curves.

The ROC curve of the algorithm is demonstrated by a particular plot (black curve) on an independent testing set for the following items:

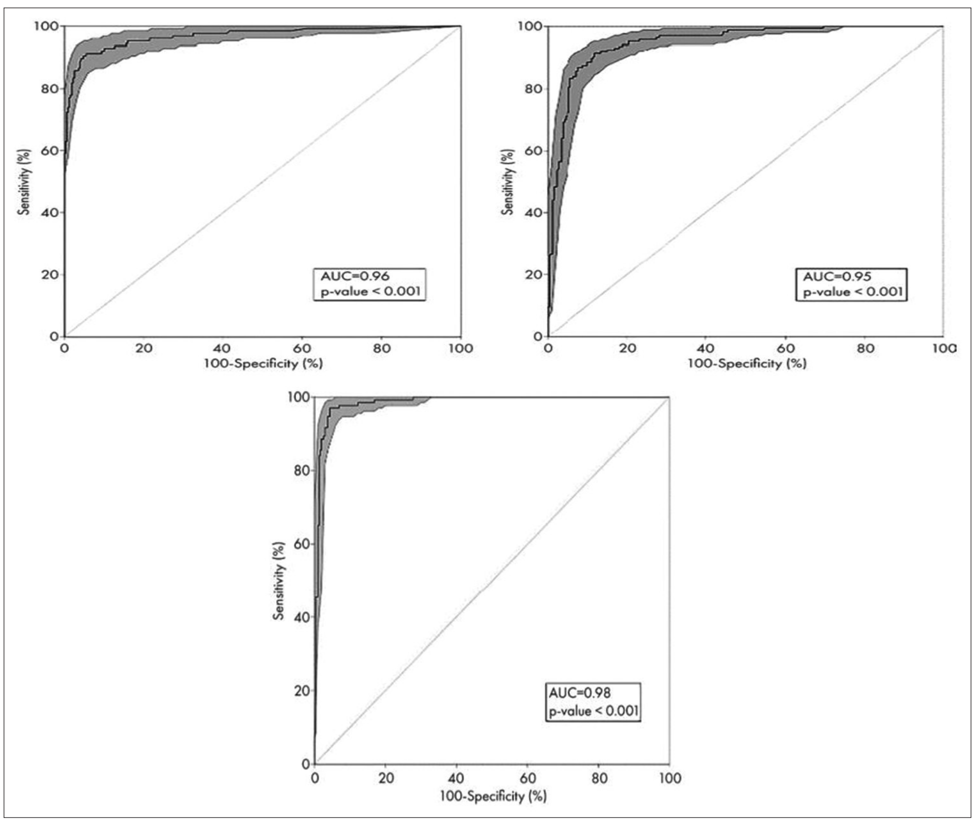

Figure 3: The receiver operating characteristic (ROC) curves model

Table 1: The data set of independent testing and the training data sets

	Testing set				Training set independent			
Parameter	CAP	COVID-19	*P*	Nonpneumonia	CAP	COVID-19	*P*	Nonpneumonia
No. of scans	690 (47)	565 (30)	…	672 (31)	165 (40)	127 (30)	…	102 (30)
No. of patients	800 (40)	200 (14)	…	632 (40)	145 (44)	60 (19)	…	120 (37)
Age (y)*	47	50	0.55	40	55	50	<0.001	41
Male patients	380 (56)	100 (53)	<0.001	287 (55)	90 (62)	35 (52)	0.17	63 (52)

Values in parentheses are percentages, CAP=community-acquired pneumonia, COVID-19=coronavirus disease 2019, *Ages are reported as means six standard deviations

(a) COVID-19: it is specified through the area under the ROC curve of 0.96,
(b) CAP: it is specified through the area under the ROC curve of 0.95, and
(c) nonpneumonia: it is specified through the area under the ROC curve of 0.98.

The gray region represents a 95% confidence level.

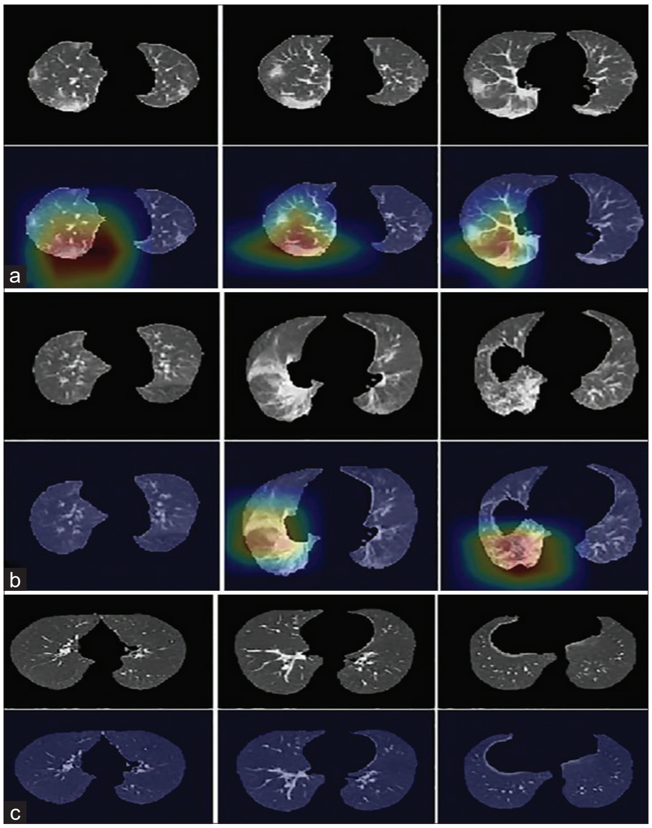

Figure 4: 3D deep learning framework for diagnosing COVID-19

The areas related to the lower parts of the ROC curve for CAP and COVID-19 are 0.95 and 0.96, respectively. To achieve proper interpretability of the model, the Gradient-weighted Class Activation Mapping (Grad-CAM) was adjusted, and visualizing the important regions was carried out by the Grad-CAM method that provides the possibility of decision making of the deep learning model. The prepared map of the localized region is created using this model without further manual interpretation. [Figure 4] demonstrates the heatmap to the suspected areas for the patients with nonpneumonia abnormalities, CAP, and COVID-19. These heatmaps show that the main focus of the prepared algorithm is on the abnormal regions while declining the regions which look normal as demonstrated in the sample with nonsevere pneumonia abnormality.

(A) COVID-19,
(B) CAP, and
(C) Nonpneumonia.

These heatmaps represent indications for standard Jet colormaps which could overlap on the original image within which the red color represents the activation area that depended on the predicted class. As could be seen from nonpneumonia abnormality example, the heatmaps show that the prepared algorithm is more sensitive to the abnormal regions while disregarding the regions that seem to be normal. Additionally, [Figure 5] represents an obvious example of a CAP case that was missorted as COVID-19. On the other hand, [Figure 6] represents an example of a case with COVID-19 which was misclassified as CAP. It should be noted that, on each row, the direction of sequential slices around the abnormality is determined to be from left to right. These cases are very challenging and to improve their accuracy, the exposure history of them should be included.

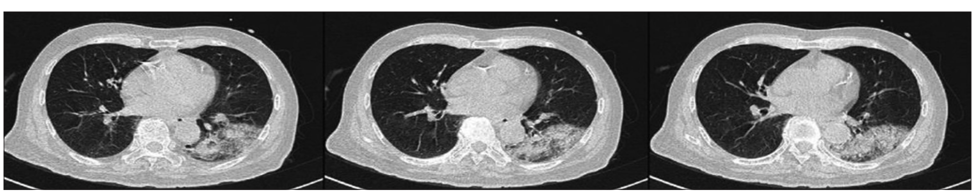

Figure 5: An obvious example of a case of community-acquired pneumonia misclassified as COVID-19. The consecutive slices around the abnormality are demonstrated from left to right

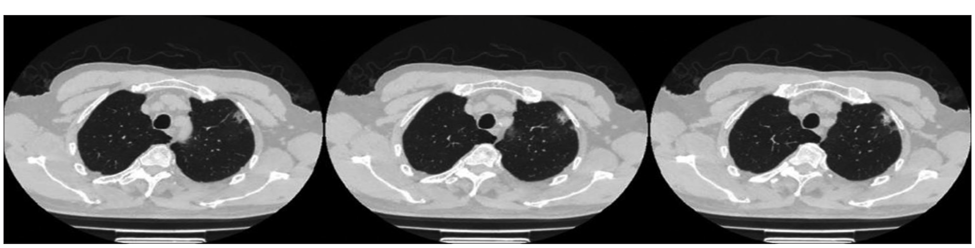

Figure 6: Representative example of case of COVID-19 that are misclassified as CAP. The consecutive slices around the abnormality are demonstrated from left to right

Table 2: The Summary of available diseases in both data sets of independent testing and training

Data set and category	No. of patients	Bacterial culture positive	Bacterial culture negative	Bacterial culture status unknown	Normal or noninfectious COVID-19
Training set					
COVID-19	200 (565 scans)	…	…	…	…
CAP	800	Bacterial pneumonia, *n*=62; bacterial and mycoplasma pneumonia, *n*=20	Viral pneumonia, *n*=14; mycoplasma pneumonia, *n*=53; pneumocystis carinii pneumonia, *n*=2	796	…
Nonpneumonia	632	…	…	…	Normal, n=239; nodule, *n*=200; COPD, *n*=102; other, *n*=11; CHF, NA; drug reaction, NA
Testing set					
COVID-19	60 (120 SCAN)	…	…	…	…
Nonpneumonia	120	…	…	…	Normal, *n*=47; nodule, *n*=42; COPD, *n*=5; other, *n*=1; CHF, NA; drug reaction, NA
CAP	145	Bacterial pneumonia, *n*=10; bacterial and mycoplasma pneumonia, *n*=1	Viral pneumonia, *n*=4; mycoplasma pneumonia, *n*=3	111	…

Discussion

In our study, the three-dimensional deep learning model was evaluated and designed to detect COVID-19 from chest CT scans. If this model be tested on an independent data set, its sensitivity and specificity in the detection of COVID-19 achieves to 90 and 96%, respectively. The areas below the ROC curves for CAP and COVID-19 were 0.95 and 0.96, respectively. Table 2 summarizes the disease statistics of patients participated in this study.

Note: Lab screening and confirmation of the disease cause was carried out on 210 patients. CAP = community-acquired pneumonia, CHF = congestive heart failure, COPD = chronic obstructive pulmonary disease, COVID-19 = coronavirus disease 2019, NA = not available, and RT-PCR = reverse-transcription polymerase chain reaction.

Here, a 3D deep learning framework was suggested to obtain a more precise diagnosis of COVID-19. The suggested framework is capable of evoking both 3D global representative and two-dimensional local features. Recent radiological studies have demonstrated that deep learning has been achieving superior performance to the other methods.[10-14] The effectiveness of deep learning techniques in detection of pneumonia on chest radiographs of pediatrics has been reported by two recent studies. Moreover, these techniques have been administered for differentiating bacterial and viral pneumonia on two-dimensional chest radiographs of pediatric patients.[13,14] In our study, 1120 CT scans of COVID-19 patients from multiple hospitals were collected. Additionally, 672 nonpneumonia and CAP CT scans were defined as the control groups for certifying the fact that detection robustness is capable of considering certain characters of equivalent imaging may be represented in patients with coronavirus in line with other kinds of lung diseases.

The main limitation of the present study is that COVID-19 is a virus disease that has similar specifications for imaging as pneumonia induced from other kinds of viruses. Anyway, because of the lack of data confirmation for lab examinations of the area for each of the cases, we were unable to pick out other viral pneumonias to compare them in the present research appropriately. Moreover, all the CAP cases from July 2016 to January 2020 were randomly selected to be included in this study. We introduce the sampling method of the present study as a sufficient one for the typical distribution of various subtypes of CAP. In this regard, any cases of organizing pneumonia, bacterial pneumonia, and non-COVID-19 viral pneumonia such as influenza virus should be included. It would be an eligible option to examine the efficiency of COVNet in differentiating coronavirus diseases from other viral infections inflaming the air sacs lungs, which have RT-PCR confirmation of the viral agent that could cause particular diseases in future studies.

Conclusion

In conclusion, the suggested pattern of robust deep learning is expanded to distinguish COVID-19 and CAP based on available chest CT scans. The presented data in this study showed that the prepared approach of machine learning applies a convolutional neural network that is capable of differentiating COVID-19 from CAP.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Yadav M, Perumal M, Srinivas M. Analysis on novel coronavirus (COVID-19) using machine learning methods. Chaos, Solitons and Fractals. 2020;139:110050. doi: 10.1016/j.chaos.2020.110050.
2. Mahdy LN, Ezzat KA, Elmousalami HH, Ella HA, Hassanien AE. Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine. medRxiv 2020. doi. org/101101/2020.03.30.20047787.
3. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 2020;121:103792.
4. Li M, Lei P, Zeng B, Li Z, Yu P, Fan B, *et al.* Coronavirus disease (COVID-19): Spectrum of CT findings and temporal progression of the disease. Acad Radiol 2020;27:603-8.
5. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, *et al.* Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. IEEE Rev Biomed Eng 2021;14:4-15.
6. Miller DD, Brown EW. Artificial intelligence in medical practice: The question to the answer? Am J Med 2018;131:129-33.
7. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018;2:719-31.
8. Yang Z, Zeng Z, Wang K, Wong SS, Liang W, Zanin M, *et al.* Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions J Thorac Dis 2020;12:165-74.
9. Albahri AS, Hamid RA, Alwan JK, Al-Qays ZT, Zaidan AA, Zaidan BB. Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): A systematic review. J Med Syst 2020;44:122.
10. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395:497–506.
11. Sethy PK, Behera SK. Detection of coronavirus disease (covid-19) based on deep features. Preprints 2020;2020030300.
12. Alom MZ, Rahman MM, Nasrin MS, Taha TM, Asari VK. COVID_MTNet: COVID-19 multi-task deep learning approaches. arXiv preprint arXiv: 2020;2004.03747.
13. Tartaglione E, Barbano CA, Berzovini C, Calandri M, Grangetto M. Unveiling COVID-19 from Chest X-ray with deep learning: A hurdles race with small data. arXiv preprint arXiv: 2020;2004.05405.
14. Alqudah AM, Qazan S, Alquran H, Qasmieh IA, Alqudah A. COVID-2019 detection using X-ray images and artificial intelligence hybrid systems 2020. doi: 10.5455/jjee. 204-158531224.
15. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. Med Image Anal. 2020;65:101794. doi: 10.1016/j.media.2020.101794.
16. Mei X, Lee HC, Diao KY, Huang M, Lin B, Liu C, *et al.* Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. Nat Med 2020;26:1224–8.