Tech Science Press

Deep Convolutional Neural Network Approach for COVID-19 Detection

Yu Xue[1,2,*], Bernard-Marie Onzo[1], Romany F. Mansour[3,4] and Shoubao Su[4]

[1]School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[2]Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing, China
[3]Department of Mathematics, Faculty of Science, New Valley University, El-Kharja, 72511, Egypt
[4]Jiangsu Key Laboratory of Data Science and Smart Software, Jinling Institute of Technology, Nanjing, 211169, China
*Corresponding Author: Yu Xue. Email: xueyu@nuist.edu.cn

Abstract: Coronavirus disease 2019 (Covid-19) is a life-threatening infectious disease caused by a newly discovered strain of the coronaviruses. As by the end of 2020, Covid-19 is still not fully understood, but like other similar viruses, the main mode of transmission or spread is believed to be through droplets from coughs and sneezes of infected persons. The accurate detection of Covid-19 cases poses some questions to scientists and physicians. The two main kinds of tests available for Covid-19 are viral tests, which tells you whether you are currently infected and antibody test, which tells if you had been infected previously. Routine Covid-19 test can take up to 2 days to complete; in reducing chances of false negative results, serial testing is used. Medical image processing by means of using Chest X-ray images and Computed Tomography (CT) can help radiologists detect the virus. This imaging approach can detect certain characteristic changes in the lung associated with Covid-19. In this paper, a deep learning model or technique based on the Convolutional Neural Network is proposed to improve the accuracy and precisely detect Covid-19 from Chest Xray scans by identifying structural abnormalities in scans or X-ray images. The entire model proposed is categorized into three stages: dataset, data pre-processing and final stage being training and classification.

Keywords: COVID-19; deep learning; convolutional neural network; X-ray

1 Introduction

The world has witnessed since December 2019, the tragic deaths of more than two million one hundred fifty-eight thousand seven hundred sixty-one (2,158,761) people out of more than hundred million two hundred thousand one hundred seven (100,200,107) registered cases of people infected with COVID-19 [1]. Caused by a virus named SARS-CoV-2 or novel Coronavirus (COVID-19) belongs to the SARS-COV family [2]. Initially registered in Wuhan, Hubei province in China, the exponential growth of positive cases worldwide in a short period of time with limited testing shows the highly contagious nature of this virus [3]. It is also noted that this infection can be transmitted to humans by vertebrates such as

bats [4]. The infection causes severe damage to the lungs, resulting in pneumonia with symptoms such as sore throat, dry cough, sneezing and high temperature [5]. In addition, some of the patients have no symptoms; the fact that they are carriers of the virus is of concern to the World Health Organization (WHO) who declared a global health emergency and coronavirus as a pandemic [6,7]. Given that a ton of effort is being made to find the effective cure for COVID-19, the main way of protection is social distancing and lockdown. In other hand, the lockdown influences the nation's GDP and has a negative psychological impact on the well-being of individuals.

Rapid detection of the virus is therefore essential to make quick decisions and to take care of patients. This detection is carried out by means of the RTPCR (Reverse Transcription Polymerase Chain Reaction) test but also by the analysis of Chest X-ray images and CT (Computed Tomography) images by a specialist. These different methods are time consuming and sometimes fail efficiently in early detection of patients. Artificial Intelligence and the Neural Networks provide many answers to help in the analysis and detection of patients with COVID-19. Indeed, several approaches have been used to address these problems, but the results show that a good margin of progress therefore remains to be made for the improvement and quality of detection to meet the crucial need for rapid and effective detachment of patients suffering from this disease.

The current way of detection of COVID-19 is the Reverse Transcription Polymerase Chain Reaction (RTPCR) test; lower and upper respiratory specimens like nasal, sputum or nasal aspirate are collected from the person suspected to be infected by COVID-19. But unfortunately, these tests could malfunction and affect the accuracy of the diagnosis, which is a major inconvenience. In addition, this testing process is very time consuming, expensive and the detection rate is also very low. Because of these problems, repeated testing must be performed to obtain an accurate diagnosis [8,9].

In this paper aim to propose a Convolutional Neural Network approach to reduce the time and effort required to perform CT scans and X-rays analysis of COVID-19-positive patients.

This paper is divided into 7 sections. The Section 2 presents the related works; the methodology used for the proposed framework is discussed in Section 3. Section 4 exposes the experiments; Section 5 shows the experimental results and comparative analysis. Conclusion and future aspects are discussed in Section 6.

2 Related Works

Studies in the field of medicine have shown that the use of X-ray images can be used to diagnose patients contaminated with COVID19 [10]. But it takes a radiologist to read and analyses these images to make decisions. On the other hand, several problems of medical science have already been tackled with the help of artificial intelligence and have shown convincing results for prediction as well as image-based disease classification [11–14]. It is therefore obvious that the use of machine learning techniques could help in the decision-making process, which is crucial and must be rapid and accurate in the context of the pandemic of COVID-19.

Many works related to the Covid-19 infected patients' detection and diagnosis using Artificial Intelligence algorithms have been proposed by researchers, by either using X-rays or CT images. Wang et al. [15] proposed a deep learning-based AI diagnostic model using CT images and obtained the accuracy of 79.3%. Xu et al. [16] proposed a model that classifies normal, COVID-19, Influenza-A viral-based pneumonia cases using CT images and records the accuracy of 86.7%. These studies use non-public datasets for testing their AI based model. Wang et al. [17] proposed an approach using public dataset of SARS-CoV-2 X-ray images and achieved an accuracy of 92.4%. Other research has been carried out with similar accuracy using X-ray images [18–21]. Another approach using deep learning

model designed with Bayes optimization was proposed by Ucar et al. [22]. To reduce data imbalance, the public dataset used were pre augmented [23].

3 Methodology

Convolutional Neural Network (CNN) is a technique that like other Deep learning techniques allows to reveal data features that are hidden from the original data. This technique has been effectively used on images and video and has led to great advances in the medical field. Unlike a simple Deep learning model, CNN has added other layers to explore in depth the features embedded in the data. They generally consist of Convolution, Activation, Pooling steps that can be repeated and intertwined, and a final classification step called Fully Connected Layer. Convolution steps involve the use of filters that are like small squares of input data that act as feature detectors from the original input and whose filters values will be learned on its own during the training process. The activation steps generally included between the convolution and pooling steps are steps where the values obtained from the convolution step are passed through an activation function to be corrected by amplifying the important elements and creates the activation map as its output. The pooling step, using a window, reduces the dimensionality of each feature map by eliminating the less useful information. Finally, the Fully Connected Layer steps are traditional Perceptron Multi-Layer or classification steps to determine the class of the input [24].

Our proposed deep learning-based COVID-19 detection comprises several phases, as illustrated in Fig. 1. The phases are summarized in the following four steps:

- Step 1: Collect the chest X-ray images COVID-19 patients and healthy persons.
- Step 2: Use data augmentation to generate 3 times more chest X-ray images per class
- Step 3: For each class divide the data into different sets: training set (81%), validation set (9%), testing set (10%)
- Step 4: Evaluate the performance of the model with: accuracy, precision, recall, F-measure.

The various parameters applied by our model for detection of COVID-19 are shown in Fig. 2.

4 Experiments

4.1 Dataset Description

For our experimentation, we mainly used the public dataset obtained from a worldwide collection by Cohen et al. [25]. This dataset consists of chest X-ray and CT images of patients which are positive or suspected of COVID-19 (Figs. 3, 4) or other viral and bacterial pneumonias (MERS, SARS, and ARDS.). Data are collected from public sources as well as through indirect collection from hospitals and physicians [26]. This dataset currently consists of around 504 X-Ray images of COVID-19 positive patients and total 866 images. The GitHub repository associated with the work collects images from websites such as radiopaedia.org, sirm.org, eurorad.org and coronacases.org. It is also open for contributions and all new images are rigorously annotated following Posterior Anterior (PA), Anterior Posterior (AP) and Anterior Posterior Supine (AP Supine) views of the lungs. Additional patient information is also provided such as: patient id, Number of days since the start of symptoms, sex, age, type of pneumonia, RT_PCR_positive, survival, if the patient was intubated , if the patient was in the ICU (intensive care unit) or CCU (critical care unit), if he needed supplemental O2, if the patient was successfully extubated, temperature, pO2 saturation, leukocyte_count, neutrophil_count, lymphocyte_count, modality (CT, X-ray, or something else), Date on which the image was acquired, 'Hospital name, city, state, country', the filename, Digital object identifier (DOI) of the research article, URL of the paper or website where the image came from, the license of the image such as

CC BY-NC-SA, 'clinical notes and other notes'). The data in image form that is used to train our model are presented in Tab. 1.
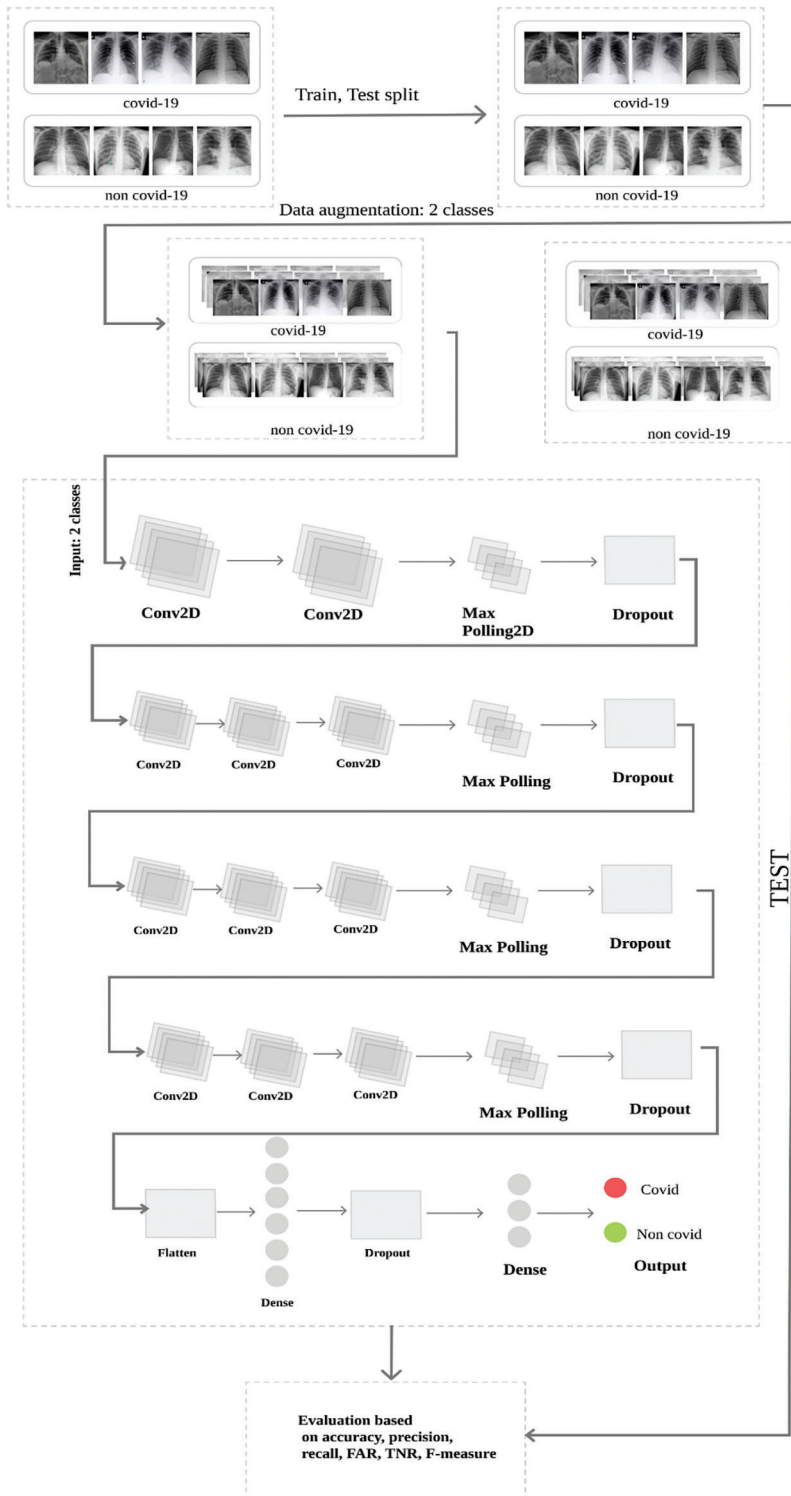

Figure 1: Proposed architecture including CNN model

```
Layer (type)                   Output Shape            Param #
=================================================================
conv2d (Conv2D)                (None, 222, 222, 64)    1792

conv2d_1 (Conv2D)              (None, 220, 220, 64)    36928

max_pooling2d (MaxPooling2D)   (None, 110, 110, 64)    0

dropout (Dropout)              (None, 110, 110, 64)    0

conv2d_2 (Conv2D)              (None, 108, 108, 256)   147712

conv2d_3 (Conv2D)              (None, 106, 106, 256)   590080

conv2d_4 (Conv2D)              (None, 104, 104, 256)   590080

max_pooling2d_1 (MaxPooling2   (None, 52, 52, 256)     0

dropout_1 (Dropout)            (None, 52, 52, 256)     0

conv2d_5 (Conv2D)              (None, 50, 50, 512)     1180160

conv2d_6 (Conv2D)              (None, 48, 48, 512)     2359808

conv2d_7 (Conv2D)              (None, 46, 46, 512)     2359808

max_pooling2d_2 (MaxPooling2   (None, 23, 23, 512)     0

dropout_2 (Dropout)            (None, 23, 23, 512)     0

conv2d_8 (Conv2D)              (None, 21, 21, 512)     2359808

conv2d_9 (Conv2D)              (None, 19, 19, 512)     2359808

conv2d_10 (Conv2D)             (None, 17, 17, 512)     2359808

max_pooling2d_3 (MaxPooling2   (None, 8, 8, 512)       0

dropout_3 (Dropout)            (None, 8, 8, 512)       0

flatten (Flatten)              (None, 32768)           0

dense (Dense)                  (None, 64)              2097216

dropout_4 (Dropout)            (None, 64)              0

dense_1 (Dense)                (None, 2)               130
=================================================================
Total params: 16,443,138
Trainable params: 16,443,138
Non-trainable params: 0
```

Figure 2: Model summary

4.2 Data Pre-Processing

In AI, to the increase in the size and diversity of labelled training sets. Data augmentation is used by generating different iterations of the samples in a data set [27]. Machine learning data augmentation is used to solve problems of class imbalance, reduce overfitting in deep learning and improve convergence [28]. For this purpose, we proceeded to the resizing of the images from $2437 \times 2806 \times 3$ in average to $224 \times 224 \times 3$ pixel images. Also, we proceeded to the flipping of images and then applied the listed augmentation technics are both original and flipped images.
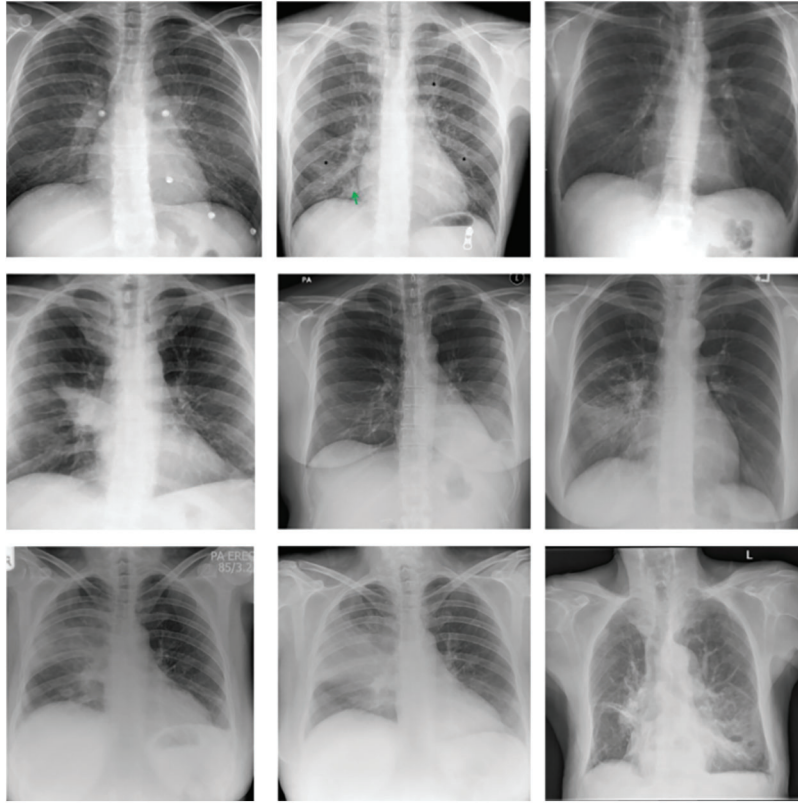
Figure 3: Images X-ray without covid-19

- Shearing
- Horizontal flip
- Zooming
- Brightness decrease
- Brightness increase

After applying augmentation, the obtained data are presented in Tab. 2.

4.3 Evaluation Metrics

The results obtained by experiments are performed system were evaluated using the evaluation criteria of precision, recall, false alarm rate, true negative rate, F-measure and accuracy. Mathematical equations of each of these measurements are respectively given in Eqs. (1)–(6).

$$\frac{TP}{TP + FP} \tag{1}$$

$$\frac{TP}{TP + FN} \tag{2}$$

$$\frac{FP}{FP + TN} \tag{3}$$

$$\frac{TN}{TN + FP} \tag{4}$$

$$2\left(\frac{Precision*Recall}{Precision + Recall}\right) \tag{5}$$
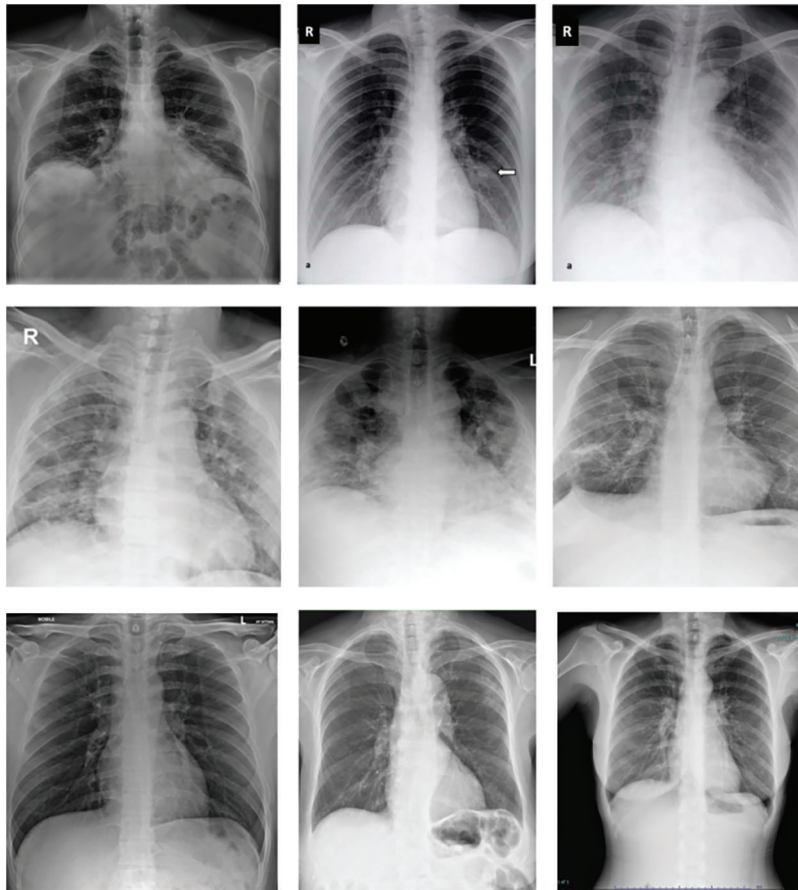
$$\frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

Figure 4: Images X-ray with covid-19

Table 1: Original dataset presentation

X-ray images	Number
Covid	504
Non_covid	362
Total	866

Table 2: Obtained dataset after applying data augmentation

X-ray images	Number
Covid	1000
Non_covid	1000
Total	2000

4.4 Environment

All the experiments are performed in python environment running on a Linux Manjaro Mikah workstation with 16Gb of RAM, a Intel® Core™ i7-8550U CPU @ 1.80GHz × 8 as processor and GeForce MX130 2048 Mb GPU with the following tools:

- ImageDataGenerator
- Keras
- Tensorflow
- Matplotlib
- Pandas
- Scikit

5 Results

Experimental results reveal that the model we are proposing has good results on training, as shown in Fig. 5. Despite the results obtained shown in the Tab. 4, we believe that improvements can be made to increase accuracy by pre-cleaning the data set. Also, the confusion matrix we're getting is in the Tab. 3. Decoding the confusion matrix, out of 30 COVID affected patients we are getting 30 people we are getting 0 wrongly classified and out of 28 normal patients we are getting 28 patients are classified right and 2 as wrongly classified.
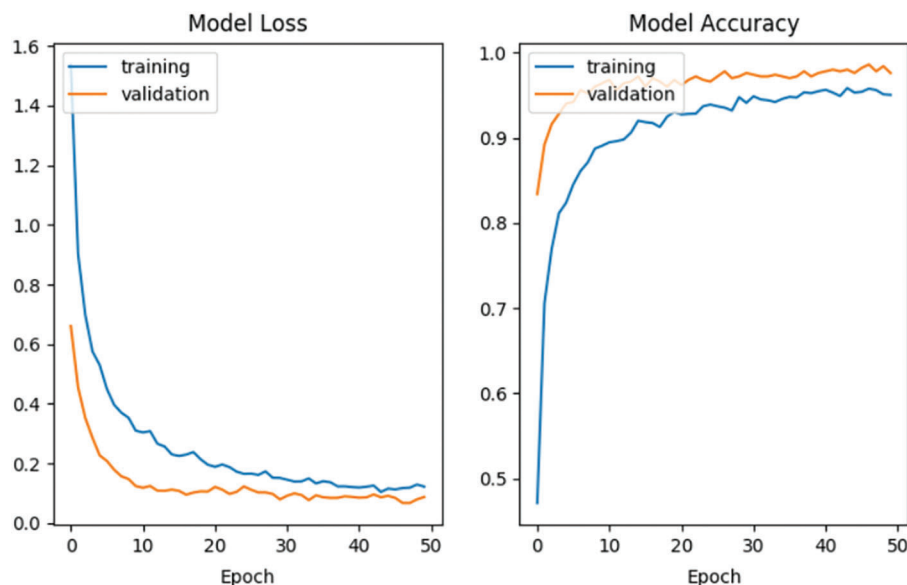

Figure 5: Loss and accuracy evolution

Table 3: Confusion matrix of the proposed

	Covid	Non-covid
Covid	TP = 30	FN = 0
Non-covid	FP = 2	TN = 28

Table 4: Covid-19 detection model metrics

Metric	Value
Precision	0.952
Recall	1
False alarm rate	0.043
True negative rate	0.956
F-measure	0.975
Accuracy	0.9767

6 Conclusion

The contribution of this study is to develop is machine-learning model for rapid and real time detection of coronavirus infected persons. This model will work as an automated tool that can be used to assist medical professional in improving the COVID-19 diagnosis accuracy. Moreover, this method could be applied to medical problems based on image analysis. Based on the experimental results, it can be concluded that the proposed model of CNN is a much better model, due to the good results (accuracy of more than 85–95%) it produces as compared to the other models. This result validates the Deep Neural Networks application for the detection of COVID-19; results that could be much improved with the use of a larger amount of data that will not need to be artificially augmented. Our future work will be in the direction of improving the Feature Selection after the convolutional layers by using Particle Swarm Optimization (PSO) on the one hand and by diversifying our data sources on the other hand.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] WHO, "WHO coronavirus (Covid-19) dashboard," 2021. [Online]. Available: https://covid19.who.int.

[2] T. T. Lam, M. H. Shum, H. Zhu, Y. Tong, X. Ni *et al.,* "Identification of 2019-nCoV related coronaviruses in Malayan pangolins in southern China," *Nature*, vol. 583, no. 7815, pp. 282–285, 2020.

[3] W. Guan, Z. Ni, Y. Hu, W. Liang, C. Ou *et al.,* "Clinical characteristics of 2019 novel coronavirus infection in China," *MedRxiv*, vol. 28, pp. 394, 2020.

[4] Y. Fan, K. Zhao, Z. L. Shi and P. Zhou, "Bat coronaviruses in China," *Viruses*, vol. 11, no. 3, pp. 210, 2019.

[5] C.-C. Lai, T. P. Shih, W. C. Ko, H. J. Tang and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (Covid-19): The epidemic and the challenges," *International Journal of Antimicrob Agents*, vol. 55, no. 3, pp. 105924, 2020.

[6] WHO, "Coronavirus disease 2019 info," 2021. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

[7] WHO, "Statement on the second meeting of the International Health Regulations, 2005 emergency committee regarding the outbreak of novel coronavirus (2019-nCoV),"2020. [Online]. Available: https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

[8] R. Liu, H. Han, F. Liu, Z. Lv, K. Wu *et al.,* "Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020," *Clinica Chimica Acta*, vol. 505, pp. 172–175, 2020.

[9] D. K. W. Chu, Y. Pan, S. M. S. Cheng, K. P. Y. Hui, P. Krishnan *et al.,* "Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia," *Clinical Chemistry*, vol. 66, no. 4, pp. 549–555, 2020.

[10] D. Wang, B. Hu and C. Hu, "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan," *China JAMA Network*, vol. 323, no. 11, pp. 1061–1069, 2020.

[11] D. S. Kermany, M. Goldbaum, W. Cai, C. C. S. Valentim, H. Liang *et al.,* "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[12] G. A. P. Singh and P. Gupta, "Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans," *Neural Computing and Applications*, vol. 31, no. 10, pp. 6863–6877, 2019.

[13] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang *et al.,* "Applications of deep learning to MRI images: A survey," *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1–18, 2018.

[14] M. K. Siddiqui, R. Morales-Menendez, X. Huang and N. Hussain, "A review of epileptic seizure detection using machine learning classifiers," *Brain Informatics*, vol. 7, no. 1, pp. 1–18, 2020.

[15] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao *et al.,* "A deep learning algorithm using CT images to screen for corona virus disease (Covid-19)," *European Radiology*, vol. 31, no. 8, pp. 6096–6104, 2021.

[16] X. Xu, X. Jiang, C. Ma, P. Du, X. Li *et al.,* "Deep learning system to screen coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.

[17] L. Wang and A. Wong, "Images, Covid-Net: A tailored deep convolutional neural network design for detection of Covid-19 cases from chest x-ray," *Scientific Reports*, vol. 10, no. 1, pp. 19549, 2020.

[18] X. Li, C. Li and D. Zhu, "Covid-Xpert: An AI powered population screening of Covid-19 cases using chest radiography images," *ArXiv*, arXiv:2004.03042, 2020.

[19] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis *et al.,* "COVID-CAPS: A capsule network-based framework for identification of Covid-19 cases from x-ray images," *Pattern Recognition Letters*, vol. 138, no. 4, pp. 638–643, 2020.

[20] M. Farooq and A. Hafeez, "Covid-ResNet: A deep learning framework for screening of Covid-19 from radiographs," *ArXiv*, arXiv:2003.14395, 2020.

[21] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir *et al.,* "Can AI help in screening viral and Covid-19 pneumonia?," *IEEE Access*, vol. 8, no. 2020, pp. 132665–132676, 2020.

[22] F. Ucar and D. Korkmaz, "COVID diagnosis-net: Deep bayes-squeezeNet based diagnosis of the virus," *Medical Hypotheses*, vol. 140, no. 2020, pp. 109761, 2020.

[23] Y. Sanajalwe, M. Anbar and S. Al-E'mari, "Covid-19 automatic detection using deep learning," *Computer Systems Science and Engineering*, vol. 39, no. 1, pp. 15–35, 2021.

[24] J. Wang, T. Zhang, Y. Cheng and N. Al-Nabhan, "Deep learning for object detection: A survey," *Computer Systems Science and Engineering*, vol. 38, no. 2, pp. 165–182, 2021.

[25] J. P. Cohen and L. D. Paul Morrison, "Covid-19 image data collection," *ArXiv*, arXiv: 2003.11597v1, 2020.

[26] D. Kermany, K. Zhang and M. Goldbaum, "Large dataset of labeled optical coherence tomography (OCT) and chest x-ray images," *Mendeley Data*, vol. 2, 2018. https://doi.org/10.17632/rscbjbr9sj.2.

[27] X. F. Li, Y. B. Zhuang and S. X. Yang, "Cloud computing for big data processing," *Intelligent Automation & Soft Computing*, vol. 23, no. 4, pp. 545–546, 2017.

[28] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin *et al.,* "Albumentations: Fast and flexible image augmentations," *Information-an International Interdisciplinary Journal*, vol. 11, no. 2, pp. 125, 2020.