



# COVID-19 Detection using Convolutional Neural Network (CNN)

Prince Chauhan<sup>1\*</sup>, Mohit Yagyik<sup>1</sup>, Shraddha Katiyar<sup>1</sup>, Praveen Bhatt<sup>1</sup>, Upasana Pandey<sup>2</sup>

## Abstract

Coronavirus disease (COVID-19) is spreading across the whole world and has habitual expressive community spread. The sudden hike in the number of patients with COVID-19, i.e. a new respiratory virus has an extreme impact on the healthcare system. There are limited kits for diagnosis, hospital beds for such kind of patients, a limited number of personal protective equipment (PPE) for healthcare temporary staff and limited ventilators. A prediction system based on deep learning can help the healthcare system to respond immediately. The X-ray images can play an important role in the early prediction of COVID-19 patients and help in the patient treatment at an earlier stage. For the disease prediction, this study presents the use of Convolutional Neural Network (CNN) that extracts the features from the images of chest X-ray. To get the edges from the images, three convolution layers with different filters are applied. Keras's Image Data Generator class is used to generate augmented images to deal with the small size of the training dataset. Classification is performed on three classes having X-ray images from COVID-19, normal people and viral pneumonia. The final results show that the proposed CNN model can predict COVID-19 patients with high accuracy.

**Keywords:** Convolution neural network (CNN), COVID-19, disease prediction, X-ray images

## INTRODUCTION

The coronavirus disease (COVID-19) declared a pandemic on March 11, 2020, by World Health Organization (WHO). It is a transmissible viral infection resulting in respiratory illness in most of the cases, proverbial to originate from Wuhan City, Hubei Province, China [1, 2]. Coronavirus disease is caused by a unique coronavirus, widely known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) previously [3]. The test is adopted as a standard diagnostic method to detect viral nucleic acid as coronavirus infection in Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) suspected individuals is called Reverse transcription-polymerase Chain Reaction (RT-PCR). RT-PCR test takes 4–6 h or maybe a whole day. A significant concern with the RT-PCR test is that it takes more time to generate results and sometimes it gives false-positive and false-negative results.

Due to the high false-negative rate of RT-PCR test and high cost and in many affected regions and countries to provide sufficient RT-PCR test kits for testing [4]. Therefore, to address these is an issue; so, chest X-ray is the commonest methodology for examining common metabolic process and respiratory organ infection for quick and efficient identification. Chest X-ray imaging is recently experimented with by varied analysis teams to detect COVID-19 by the radiographic images of patients. COVID-19 detection from chest X-ray rules out the prospect of the doctor obtaining infection by the COVID-19 patients [5]. The foremost challenge is the similarity between

### \*Author for Correspondence

Prince Chauhan  
E-mail: prince.chauhan.90038@gmail.com

<sup>1</sup>Student, Department of Information Technology, IMS Engineering, Ghaziabad, Uttar Pradesh, India

<sup>2</sup>Associate Professor, Department of Information Technology, IMS Engineering College, Ghaziabad, Uttar Pradesh, India

Received Date: July 12, 2021

Accepted Date: August 25, 2021

Published Date: September 25, 2021

**Citation:** Prince Chauhan, Mohit Yagyik, Shraddha Katiyar, Praveen Bhatt, Upasana Pandey. COVID-19 Detection using Convolutional Neural Network (CNN). Journal of Artificial Intelligence Research & Advances. 2021; 8(2): 22–30p.

---

COVID-19 and pneumonia caused by alternative infections [4]. In this research work, the Convolutional Neural Network (CNN) classifier is used to separate COVID-19 cases from normal and pneumonia cases [6].

## RELATED WORK

In the paper, "CovidAID" [3] the authors worked on the chest X-ray which helpful for the system to make out the decision that whether the patient would be live with the other patient or isolate within a ward and that also helpful for the false-negative RT-PCR. In this model, CheXNet is used with an accurate measure of 90.6% with 100% recall for COVID-19 detection.

In the paper "COVID-19 Screening" [7], the authors worked on X-ray images because the CT scan images are not much enough to provide good results; so, the X-rays were used to detect the impact of COVID-19 on an individual patient. Here, 18 layers of residual convolution layers were used, pre-trained on the ImageNet dataset. In the training phase, that produced a sensitivity rate of 90% and specificity of 87.84% with the X-ray dataset. But the limitation is that it misses approximately 4% of corona cases and about 30% in the error in positive rate [8].

The paper is "Automatic Detection of Coronavirus Disease" [9]. Here different classification models like ResNet50, ResNet101, ResNet152, InceptionV3 are used to detect COVID-19 by the chest X-rays image belonging to different three classes; the learning algorithm is used instead of ImageNet because it takes inadequate data and training time. From all the above classifiers, the ResNet50 and ResNet101 have an overall output of 96.2% with nearly 99% to detect the COVID-19. The highest accuracy for the three different datasets for the COVID-19 detection is for Dataset 1: accuracy is 96%, Dataset 2: accuracy is 99.5%, and for Dataset 3: accuracy is 99.7%.

The Paper is "Transfer Learning with Deep Convolutional Neural Network for Pneumonia Detection" [10]. Here, MATLAB was used in training and examine the model, in these image sets are going through some pre-processing steps. Dataset Augmentation and training are done by pre-trained algorithms like AlexNet, ResNet18, DenseNet201 and SqueezetNet. From the above all the pre-trained algorithms, the highest accuracy is given by DenseNet201 which is 95%.

The paper is for "COVID-19 Detection" so Here, there are two models are used that are ResNet-50 and VGG-16 and this paper also describes a CNN model for the detection of COVID-19 by the use of multiple image datasets; in this it takes the suspects of CT scan and suspect of chest X-rays of the individuals. For feature extraction, the proposed model does not need anything. This model gives the highest accuracy for the CT scan images that is approximately 95.39% and the accuracy for the X-ray images is about 98.98%.

## MATERIAL AND METHODS

### Dataset

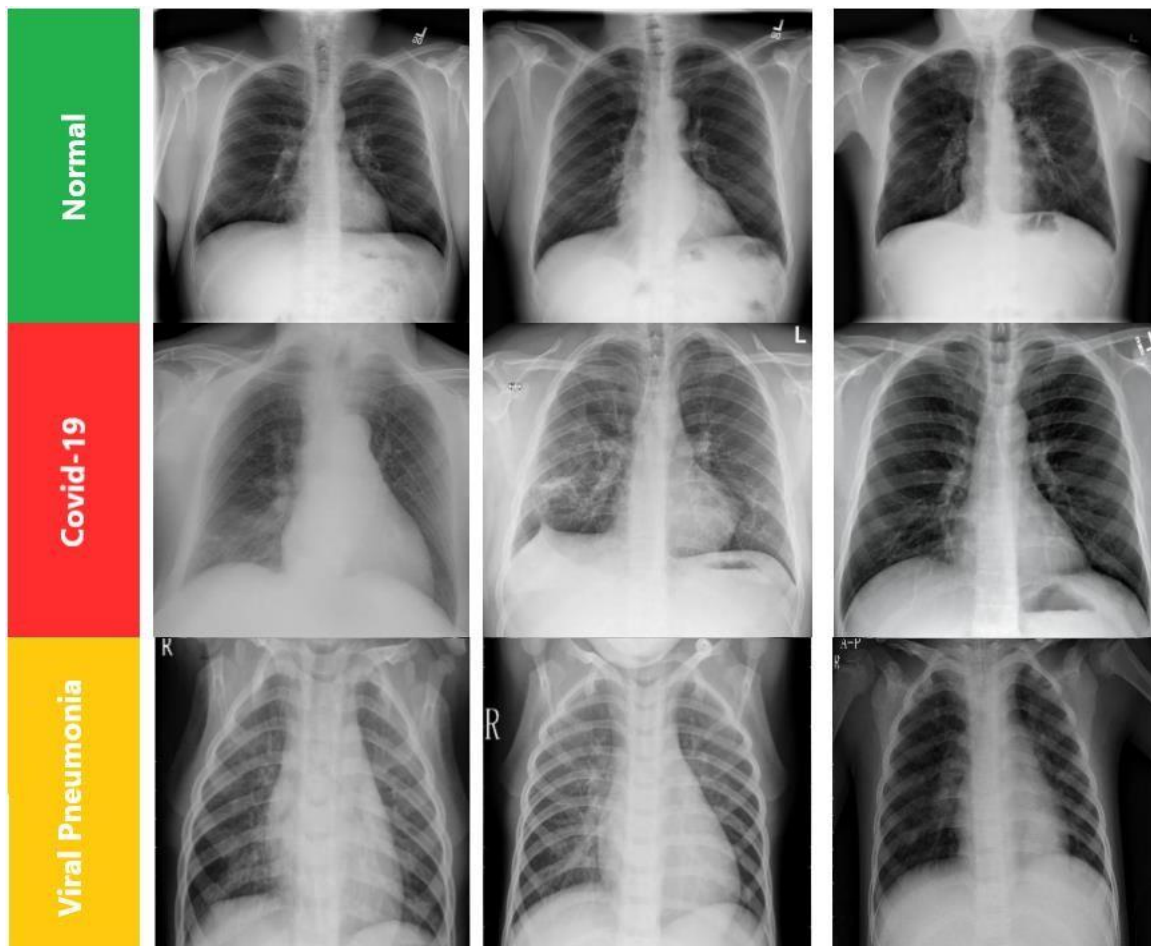
In this study, we have used 3829 chest X-ray images from the Kaggle repository, called "COVID-19 Chest X-Ray Database" [10]. These images are divided into three classes: COVID-19, Normal and Viral Pneumonia. Further, these images are divided into training, testing and validation classes.

Our proposed model is based on three binary created datasets (Training, Testing, and Validation) with chest X-ray images.

In the pre-processing phase, all images are resized to 256×256 pixels in the dataset. As shown in Table 1, the dataset consists of 1143 X-ray images of COVID-19, for normal it is 1341 and the number of infected people by viral pneumonia are 1345. Figure 1, represents Kaggle's repository of chest X- ray images of viral pneumonia, COVID-19, and normal patients.

**Table 1.** Number of chest X-ray images per each class.

Classes	COVID-19	Normal	Pneumonia	Total
Training	878	1076	1080	3034
Testing	255	255	255	765
Validation	10	10	10	30
Total	1143	1341	1345	3829



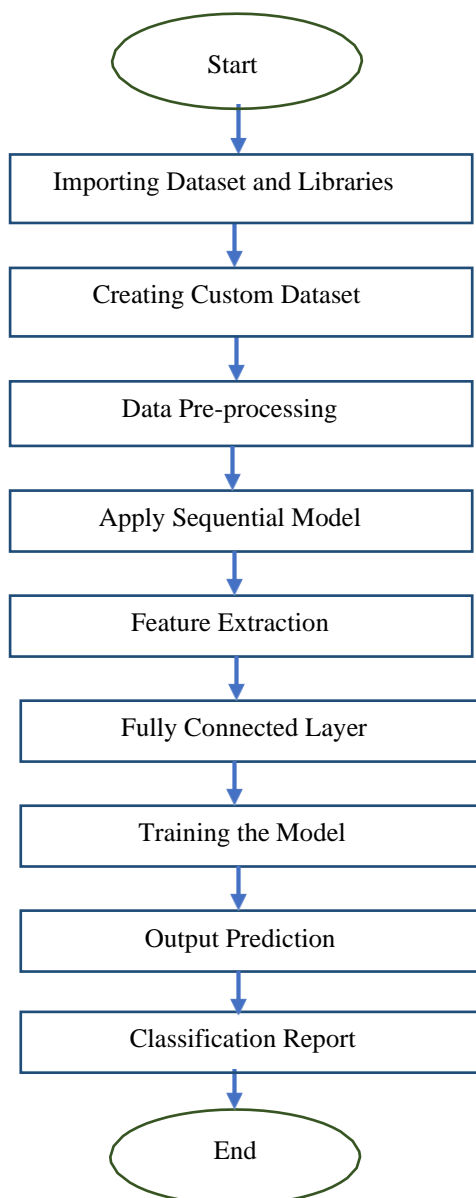
**Figure 1.** Chest X-ray images of viral pneumonia, COVID-19, and normal patients.

### PROPOSED MODEL DESCRIPTION

In this study we proposed a new model of deep learning for predicting the status of coronavirus in the chest X-ray images of the patients. Figure 2 shows the complete workflow of the proposed model for classifying the viral pneumonia or COVID-19 status in chest X-Ray images of the patients. Our proposed model is created on a convolutional neural network (CNN).

For image recognition, convolutional neural network (CNN), a class of deep neural networks is used [11]. In CNN, the input image is converted into a format that can be processed by the computer. Because of that reason, images are firstly converted to matrix format. And then the system determines which image belongs to which category based on the differences in images. The model learns these differences on the label during the training phase. And then trained model makes predictions for input images.

Our proposed model comprises of three main phases to achieve the analytic procedure of COVID-19 virus, these steps are as follows:



**Figure 2.** Workflow of proposed model for categorizing the Pneumonia COVID-19 status in chest X-Ray images.

### Step 1: Pre-processing

All chest X-ray images are stored in one dataset (dataset is available on Kaggle [10]) and provided for rescaling at a same size of  $256 \times 256$  pixels to be appropriate for feature extraction.

### Step 2: Training the Model and Validation

To start this phase of proposed deep learning models, according to the Pareto principle, the pre-processed dataset is split in the ratio of 80-20. It means 20% of X-ray images are used for testing purpose. Also, 80% chest X-ray images are used for training our proposed model.

### Step 3: Classification

At the end of our proposed model, the testing images are served to the given deep learning-based classifier to classify all the chest X-ray image batches into one of three classes: positive viral pneumonia, normal case or COVID-19. In the end, to analyze the overall performance of the model, various performance matrices were calculated based on the following metrics described:

### Classification Performance Analysis

To calculate the performance of our proposed model, there are different metrics applied in this study. First, the confusion matrix [12] is created using the cross-validation estimator [13], as shown in Table 2. A confusion matrix can have four predicted results listed as follows:

- **True Positive (TP):** It means correctly diagnosed as diseased (For e.g., COVID-19 is predicted as COVID-19 positive case).
- **True Negative (TN):** In this the model correctly forecasts the negative class (For e.g., Normal image is predicted as normal).
- **False Positive (FP):** It means incorrectly diagnosed as diseased (For e.g., COVID-19 is predicted as viral pneumonia).
- **False Negative (FN):** In this the model incorrectly predicts the negative class (For e.g., Normal image is predicted as COVID-19).

**Table 2.** Confusion matrix.

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

The following performance metrics can be measured by computing the values of probable results in the given confusion matrix.

#### Accuracy

Accuracy is considered as the most important metric to evaluate performance of deep learning classifier [14], as shown in the formula. It can be calculated as adding the true negatives forecast and true positives forecast divided by the adding total component values of the given confusion matrix.

$$Accuracy(\%) = \frac{TP + TN}{TP + FP + FN + TN} 100\%$$

#### Precision

It can be calculated by dividing true positive forecast by summation forecast of false positive and true positive.

$$Precision = \frac{TP}{TP + FP}$$

#### Recall

It can be calculated by dividing true positive forecast by summation forecast of false negative and true positive.

$$Recall = \frac{TP}{TP + FN}$$

#### F1-score

It estimates overall accuracy of the model that integrates recall and precision (given in formula). It is the double of the proportion between the multiplication of recall and precision, divided by the sum of recall and precision metrics.

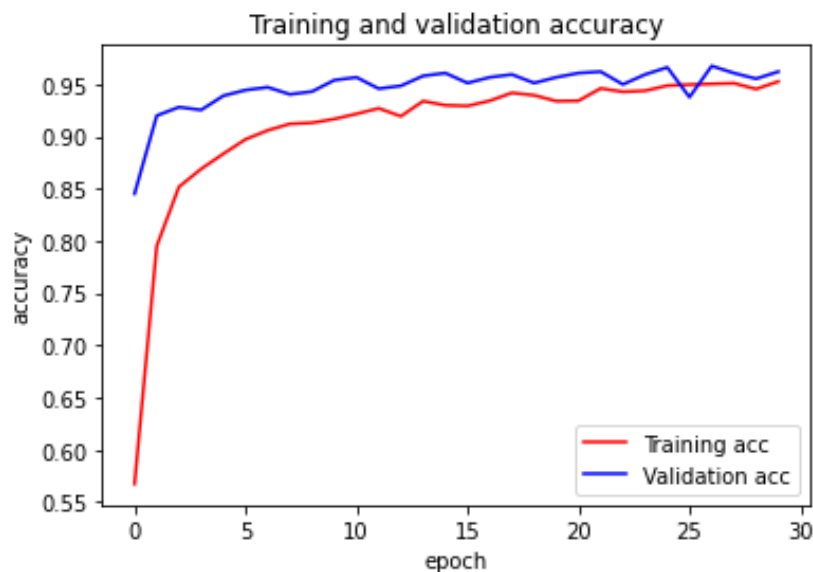
$$F1\ score = 2 \left( \frac{Precision * Recall}{Precision + Recall} \right)$$

## RESULTS

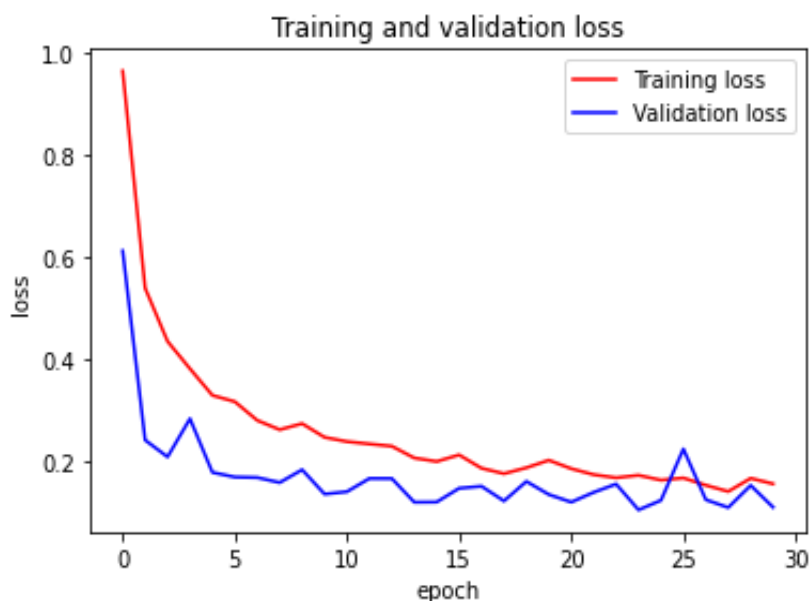
In this study, we train the proposed model with three different classes (Normal, Viral pneumonia and COVID-19). We made a custom dataset having 20% of the data kept for testing and the remaining 80% was kept for training. This process was resumed until every 20% part of the dataset was tested by the model.

Figure 3 represents the accuracy of training and test dataset images. As it is clear from the graph, we have achieved 94.74% accuracy for the training dataset and 96.20% accuracy for test dataset images.

Figure 4 represents the value of loss while training and testing the images. As shown in Figure 4, the loss for training dataset images was 0.1623 and for the test dataset, it was 0.1253.



**Figure 3.** Training and validation accuracy graph.



**Figure 4.** Training and validation loss graph.

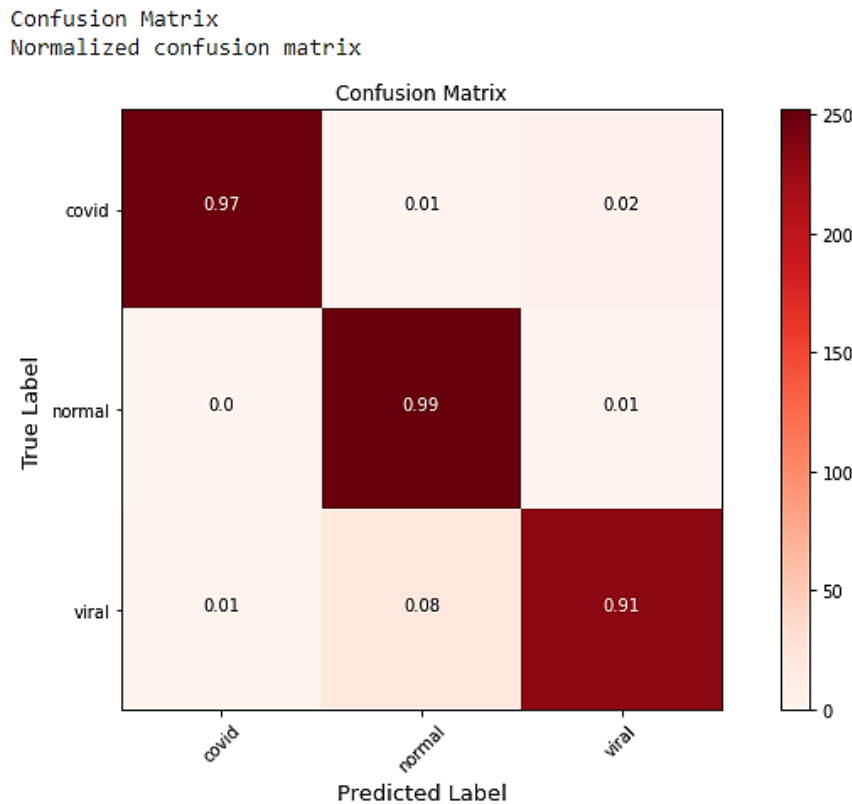
Figure 5 shows final prediction of COVID-19 (first image), normal (second image) and viral pneumonia (third image) patients. The prediction is labelled with green text because that was correctly predicted. In case of the wrong prediction, the output will be labelled with red colour.

In Figure 6 the normalized confusion matrix shows us that our model has achieved 96% accuracy for the test dataset. Viral pneumonia is predicted as normal because this class has 91% recall power.

Table 3 shows the classification report for the test dataset. As shown in Table 3, we have obtained an overall 96% accuracy with 765 X-ray images (255 X-ray images of the chest in each class). In the case of COVID-19, the proposed model achieves 99% precision with 97% sensitivity (recall).



**Figure 5.** Final prediction of COVID-19 (first image), normal (second image) and viral pneumonia (third image) patients.



**Figure 6.** Confusion matrix.

**Table 3.** Classification report.

	Precision	Recall	F1-score	Support
COVID-19	0.99	0.97	0.98	255
Normal	0.92	0.99	0.95	255
Viral	0.97	0.91	0.94	255
Accuracy			0.96	765
Macro avg.	0.96	0.96	0.96	765
Weighted avg.	0.96	0.96	0.96	765

## DISCUSSION

There are many methods to detect coronavirus in patients, but deep learning-based systems are very common in this pandemic. You can find much research on COVID-19 detection, some of these are given in related work. For two classes, the main goal is to differentiate between COVID-19 negative cases from COVID-19 positive cases. For three classes it is very hard to differentiate between viral pneumonia and COVID-19 patients. Viral pneumonia is another type of diseases infecting the lungs of the patient.

Arpan Mangal did the research for three types of classes (Normal, Viral Pneumonia and COVID-19) and obtained 90.6% accuracy with 100% sensitivity (recall) for the COVID-19 infection [1].

Jianpeng Zhang [2] conducted studies and obtained a 90.00% sensitivity and 87.84% specificity. In this study, the dataset contains 70 X-ray images of COVID-19 infected patients and 1008 X-ray images of viral pneumonia infected patients.

In this study, we have obtained an overall 96% accuracy with 3829 X-ray images. For the COVID-19 class, our proposed model achieves 99% precision with 97% sensitivity (recall). It shows our

model is better than Arpan Mangal and Jianpeng Zhang studies. The performance of our model in the COVID-19 class is significantly higher than the other two classes (normal and viral pneumonia). In this study, we have used more chest X-ray images than many studies.

## CONCLUSION

Finally, we have proposed a deep learning network that will identify the normal, COVID-19 and viral pneumonia patients with an accuracy of 96%. This model is trained on 3034 images and at the same time tested on 765 images where 255 images belong to each class. In the end, we are validating our model on 30 images. This network is useful for detecting COVID-19 at an early stage. In future, we want to add some more images to our dataset to improve accuracy.

## ACKNOWLEDGEMENT

We are deeply thankful to Upasana Pandey (Associate Professor of IT department); without her constructive feedback, this study would not have been a success. The valuable advice and suggestions for the modifications, corrections and improvement did enhance the perfection in the study.

## REFERENCES

1. Roosa K, Lee Y, *et al.* Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect Dis Model.* 2020; 5: 256–263.
2. Li Yan, Hai-Tao Zhang, *et al.* Prediction of criticality in patients with severe COVID-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. *medRxiv.* 2020.
3. Simondon A, Rolland P, Mailles A, Campese C, Stoecklin SB, Silue Y, *et al.* First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. *Euro Surveill.* 2020; 25(6): 2000094.
4. Amit Kumar Jaiswal, Prayag Tiwari, *et al.* Identifying pneumonia in chest X-rays: A deep learning approach. *Measurement.* 2019; 145: 511–518.
5. Arpan Mangal, Surya Kalia, *et al.* CovidAID: COVID-19 Detection Using Chest X-Ray. *Electrical Engineering and Systems Science.* 2020.
6. Jianpeng Zhang, Yutong Xie, *et al.* Viral Pneumonia Screening on Chest X-ray Images Using Confidence-Aware Anomaly Detection. *Electrical Engineering and Systems Science.* 2020.
7. Novitasari Dian Candra Rini, Hendradi Rimuljo, *et al.* Detection of COVID-19 chest x-ray using support vector machine and convolutional neural network. *Commun Math Biol Neurosci.* 2020; 706499: 1–19.
8. Tawsifur Rahman, Chowdhury Muhammad EH, *et al.* Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray. *Appl Sci.* 2020; 10(9): 3233.
9. Dakshina Ranjan Kisku, Mamata Dalui, *et al.* COVID-19 Detection on Chest X-Ray and CT-Scan Images Using Multi-image Augmented Deep Learning Model. *Biorxiv.* 2020.
10. Kaggle. COVID-19 Radiography Database: COVID-19 Chest X-ray Database. [Online]. Available from <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.
11. Nadia Jmour, Sehla Zayen, *et al.* Convolutional neural networks for image classification. 2018 International Conference on Advanced Systems and Electric Technologies (IC\_ASET); 2018 Mar 22–25; Hammamet, Tunisia. New York: IEEE; 2018.
12. Md Zahangir Alom, Shaifur Rahman MM, *et al.* COVID\_MTNet: COVID-19 Detection with Multi-Task Deep Learning Approaches. *Electrical Engineering and Systems Science.* 2020.
13. Scikit Learn. `sklearn.metrics.confusion_matrix` [Online]. Available from [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion\\_matrix.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.confusion_matrix.html).
14. Nirschl Jeffrey J, Andrew Janowczyk, *et al.* A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of H&E tissue. *Plos One.* 2018 Apr; 13(4): e0192726. [Online]. Available from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0192726>.