Detection of COVID-19 and Classification of Pneumonia Using Deep Neural Networks Algorithms

G. Bhavitha, A. Harsha Vardhan, K. Nikhil, Sk. Tasleema Farhan, and V. Saritha

Abstract COVID-19 which is a subclass of severe acute respiratory syndrome (SARS) is a viral disease which emerged from China in 2019. At first, there are shorthand of test kits available to diagnose the COVID-19 disease. The tests available to diagnose the COVID-19 are RT-PCR (real-time polymerase chain reaction), Rapid Antigen test and Antibody test. But in these, only RT-PCR has the high accuracy, and it is a time-taking process. It takes nearly from 4 to 48 h. Here, AI plays an important role in diagnosing the disease. In the recent years, AI becomes a part of medical field and is widely used in classification. The chest X-Rays are used to detect the COVID-19 using deep learning and the model used to detect the COVID-19 is ResNet18 which is a residual network containing 18 layers. In this work, we classified four types of classes to make sure that our model performance is better and classify accurately. The data set contain a total of 5365 images. In this, we used 80% of data for the training and 20% for validation. The accuracy obtained in classification of three classes is 96.67% and for four classes, the accuracy is 91%. We have also used another model for comparison which ResNet50 and achieved an accuracy of 75%.

Keywords Deep learning · Resnet18 · Resnet50 · COVID-19 · Python

1 Introduction

COVID-19(SARS COV2) is a form of pneumonia. First emerged at Wuhan Seafood Wholesale Market (Huanan Market) in Wuhan, China. The cases were identified from 2019. It is a positive-stranded RNA virus belongs to Coronaviridae, a subgroup B beta corona virus. It is also a 96.2% sequence of RATG13 virus. It infects endothelial cells living in blood vessels. The symptoms are fever, cough, short breathing, fatigue, headache, smell lose, sour throat and running nose. The total confirmed cases are 1165,772,430 and the total deaths are 3,437,545 as per 22 which is worldwide.

G. Bhavitha (✉) · A. Harsha Vardhan · K. Nikhil · Sk. Tasleema Farhan · V. Saritha
Department of Electronics and Communication Engineering, VR Siddhartha Engineering College (Autonomous), JNTU, Kakinada, India

linear   logarithmic
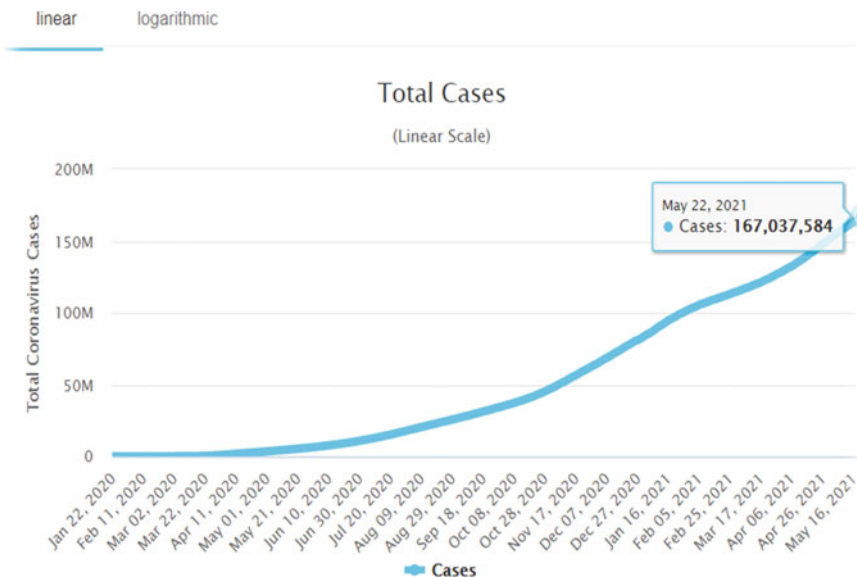
Total Cases

(Linear Scale)

Fig. 1 The total corona cases worldwide

The most affected countries are the USA, India and Brazil respectively. The initial days, there are no test kits available readily to test the patients. The tests used to detect the corona virus are RT-PCR test, Antigen test and Antibody test. RT-PCR test is a gene detection method. It collects samples from nasal and throat. It requires high laboratory and takes 48 h to know whether a person is infected with COVID or not. The accuracy of this test is 97%. Antibody test takes blood samples to detect the antibodies against the COVID-19. It is fast but the accuracy is low and can be misleading. Rapid antigen test takes samples from nose to test the viral proteins. It takes 15 min and is not recommended for COVID detection. So as the RT-PCR test takes nearly 48 h which is time-consuming process, deep learning comes in place to support the doctors in detection of COVID-19 by performing convolution neural networks on x-rays. Nowadays, artificial intelligence produces a stable and accurate results, and it also plays an important role in the health industries. There are many techniques used to detect the COVID-19 such as using networks likeVGGNet16, Alex Net, InceptionV3 and MobileNetV2. In this work, we are using ResNet50 for classification of pneumonia and corona virus (Figs. 1 and 2).

2   Related Works

In paper [1], Amit Kumar has taken a total of 771 images. They have used various neural networks such as DenseNet201, ResNet50 and InceptionV3 and achieved an
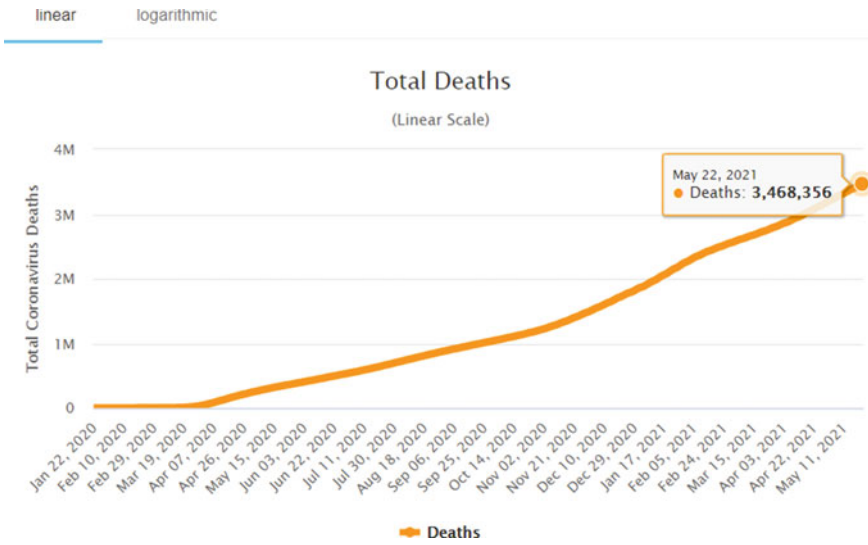
Fig. 2 The total deaths worldwide

accuracy of 93.6%, 95.3%, 94%, respectively, and sensitivity of 92%, 98% and 93%, respectively. They classified two types which are COVID +ve and COVID −ve. The total images for training are 771 images and for testing, the images are 235. They have used a sigmoid function in fully connected layer.

In paper [2], Linda Wang proposed a model of COVID-Net architecture. This is an open source available in [3]. The model was pretrained on the ImageNet model and achieved an accuracy of 93.3% and sensitivity of 95%. In this model, the classification is multi-classification and the classified types are normal, non-COVID-19 pneumonia and COVID-19. The data set contains a total of 1975 images. The data source is available here [3].

In paper [4], Afshar has proposed a model of COVID-caps architecture. This model has four convolutional layers and three capsule layers. This model classifies COVID +ve and COVID –ve with an accuracy of 98.3%. The fully connected layer uses sigmoid function.

In paper [5], Ahishali proposed a Chex Net architecture. This model has DenseNet-121 and a COVID-19Net architecture. It uses SoftMax function for multi-classification in output layer. This model classifies normal, viral pneumonia, bacterial pneumonia and COVID-19. The data set used is COVID Family data set. The accuracy obtained is 88.9%.

In paper [6], Qin used a CovNet model. This model is a 3D deep learning model. This model uses ResNet as a base model. The SoftMax function is used in fully connected layer. This model classifies a multi-classification. The classified types are COVID-19, CAP and non-COVID-19 pneumonia. The accuracy achieved is 90%.

In paper [7], Rachna made a comparison between different neural networks. The model used in this paper are Xception, InceptionV3 and ResNeXt. The accuracy

achieved by this model are 97.97%, 96% and 93%, respectively. The total images used for classification is 6432 images. In these, 5467 images are used for training and 965 images for validation. The classification is multi-classification and the classified types are COVID-19, pneumonia and normal. The data set can be accessed here [8].

In paper [9], a deep learning convolution neural network with seven layers is used. The results of image processing and feature analysing COVID-19 gives 91.67% accuracy. Again, the data is further classified and trained using CNN for feature extraction. It achieves 100% accuracy.

3   Methods and Materials

3.1   *Proposed Approach*

The following model shows the proposed approach in this paper. The steps included in this model are Data Collection, Data Pre-processing, Training the model and results analysis (Fig. 3).
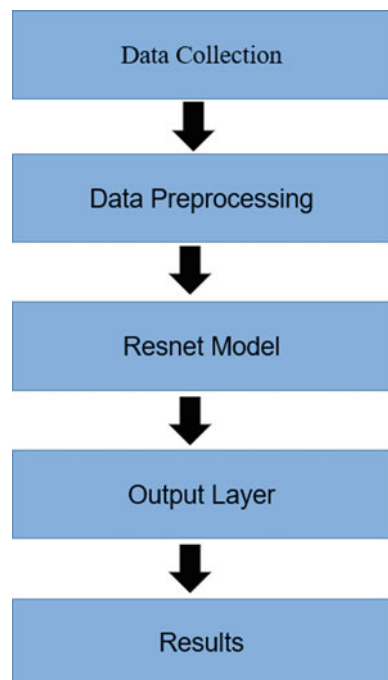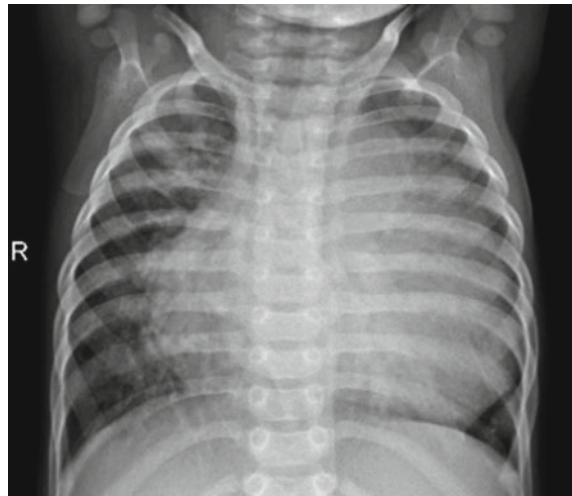
Fig. 3 Proposed approach

Table 1 Data set distribution

Label	Training samples	Testing samples
COVID-19	1200	268
Normal	1341	268
Viral pneumonia	1345	268
Bacterial pneumonia	1478	268

**Fig. 4** Bacterial pneumonia

3.2 Data sets

The data sets used for this model are collected from Kaggle repository. The data sets are openly accessible to everyone and can be accessed here [10]. The bacterial data set can be accessed here [11]. A total of 5365 images have taken for model validation. The data set contains 1341 normal images, 1345 viral images, 1200 COVID images and 1478 bacterial images. The data set distribution is shown in Table 1.

Some of the examples of chest X-rays are: (Figs. 4, 5, 6 and 7)

3.3 Data Pre-processing

All the images in the data set have undergone by the following steps: (10 The images are reshaped into (224,224,3) (2) The flip used here is RandomHorizontalFlip. (3) The Zoom range is 0.5 and Sheer range is 0.5. The framework used here is Py torch.

Then, the images are applied to the trained model. By using Random_Split method, the data set is divided into 80% for training purpose and 20% for validation.

Fig. 5 Normal
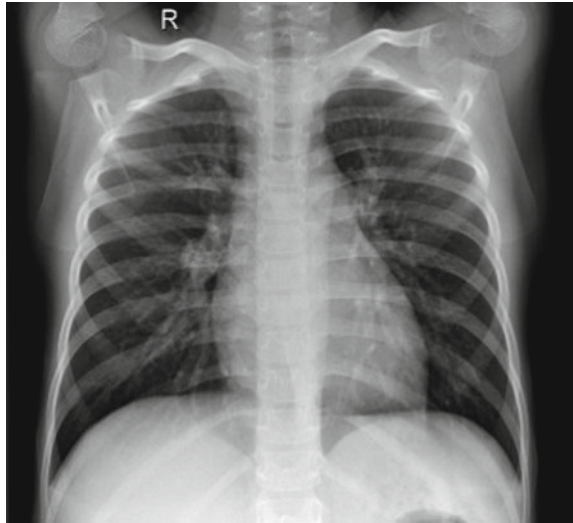
Fig. 6 COVID-19

3.4 Proposed Neural Networks

3.4.1 ResNet18

ResNet18 is a residual network having 18 layers. This model has identity connection block and convolution block. The convolution block will be applied if the input size
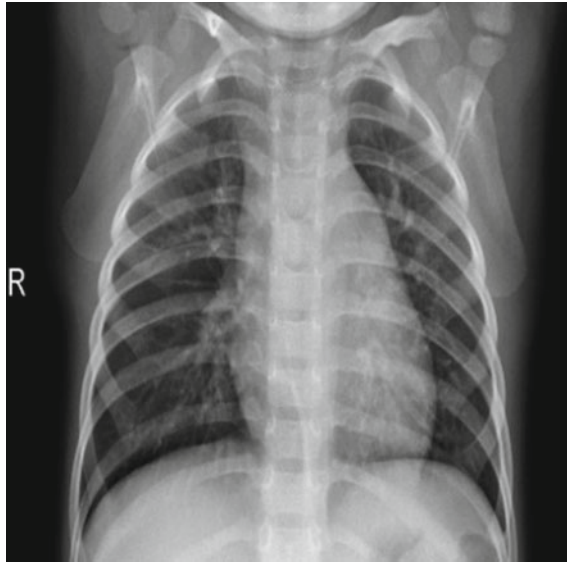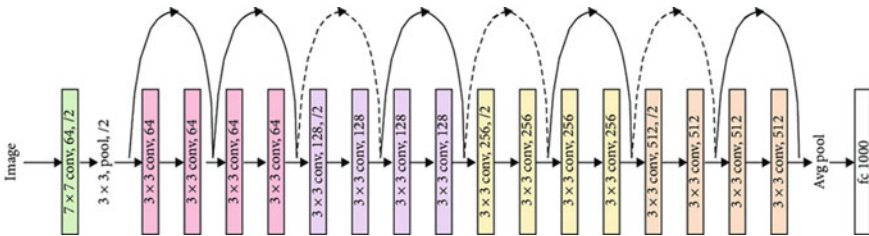
Fig. 7 Viral pneumonia

Fig. 8 Resnet-18 architecture

and output size are not matched and identity connection block is applied if the input and output sizes are matched. The loss function used in this model is categorical entropy loss, and the SoftMax function is used since it is a multi-classification. The total trainable parameters are 11,178,564. The model used here is pre-trained model. The model architecture is shown below.

To train this model Google Colab and Pytorch framework is used (Figs. 8 and 9).

3.4.2 ResNet50

ResNet50 is a residual Network having 50 layers. This model has identity connection block and convolution block. The convolution block will be applied if the input size and output size are not matched and identity connection block is applied if the input
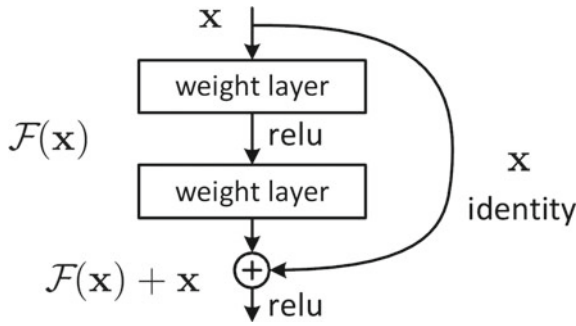
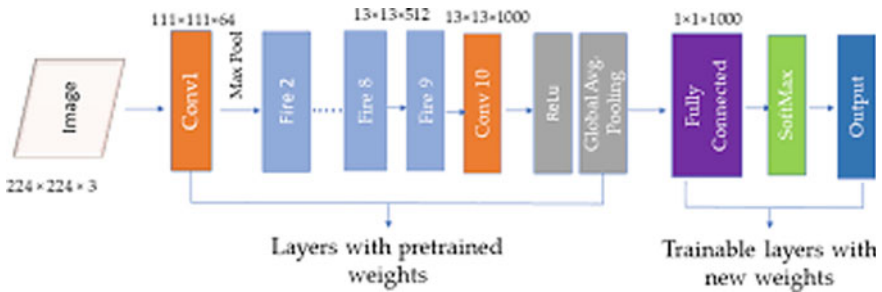**Fig. 9** Basic building block of residual network

Fig. 10 Resnet50 architecture

and output sizes are matched. The loss function used in this model is categorical entropy loss, and the SoftMax function is used since it is a multi-classification. The total trainable parameters are 31,956,868. The model used here is pretrained model. The model architecture is shown in Fig. 10.

To train this model, Google Colab and Keras Framework is used.

4 Result Analysis

The proposed models are evaluated using confusion matrix and by using following parameters such as sensitivity, specificity, accuracy, $F1$ score and precision.

$$Accuracy = (TP + TN) * 100/(TP + TN + FP + FN)$$

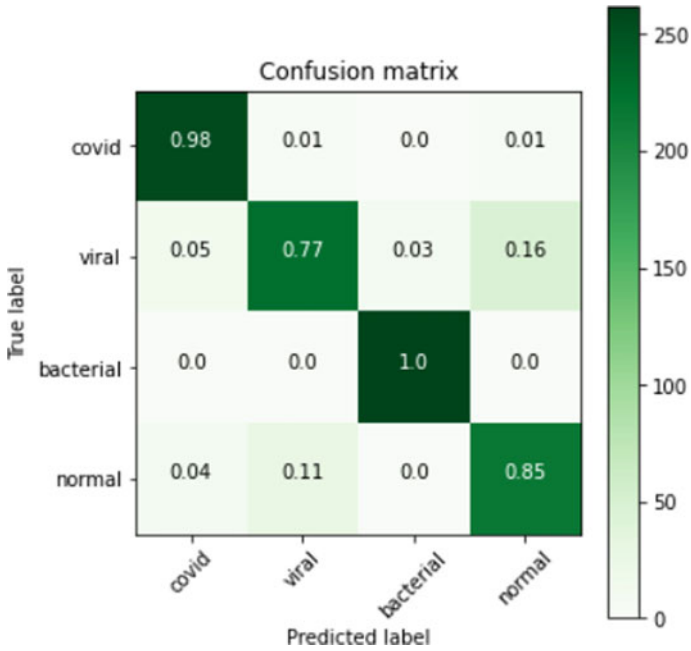$$Sensitivity \, or \, Recall = TP/(TP + FN)$$

$$Precision = TP/(TP + FP)$$

Fig. 11 Confusion matrix with normalization

$$\text{Specificity} = TN/(TN + FP)$$

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision})/(\text{Recall} + \text{Precision})$$

4.1 ResNet-18 Result Analysis—Confusion Matrix

To get the confusion matrix and classification report import confusion matrix from sklearn (Figs. 11 and 12).

Classification Report

See Table 2.

4.2 Resnet-50 Result Analysis—Confusion Matrix
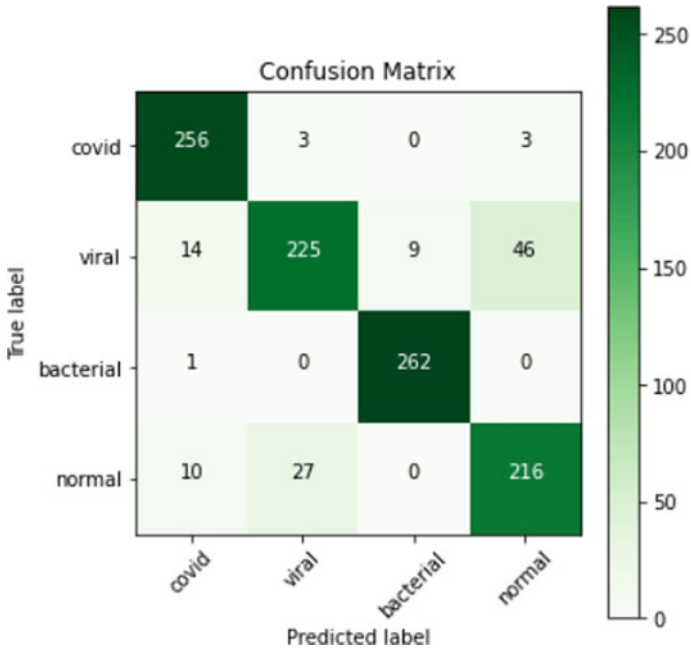
See Figs. 13 and 14.

Classification Report

Fig. 12 Confusion matrix without normalization

Table 2 Testing data report

Label	Precision	Recall	F1-score	Support
COVID	0.95	0.98	0.96	266
Viral	0.85	0.82	0.83	273
Bacterial	1	0.97	0.98	256
Normal	0.84	0.86	0.85	277

See Table 3.

4.3   Comparison Between Two Models

See Table 4.

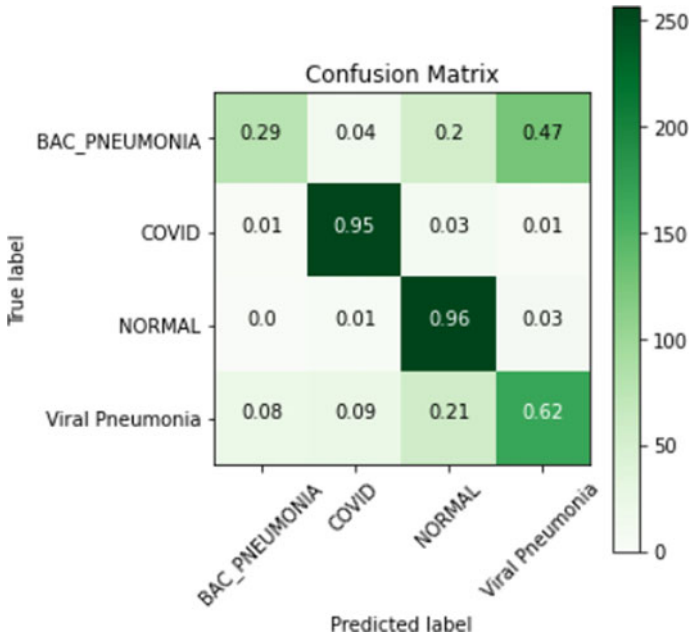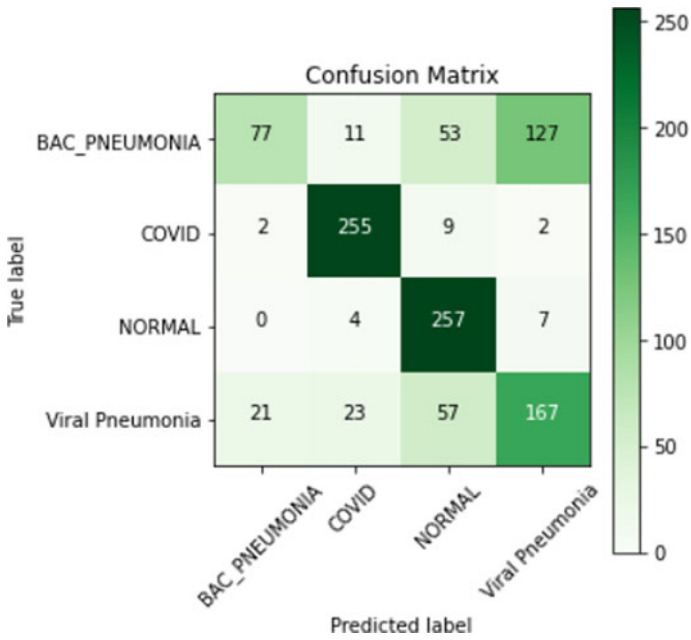**Fig. 13** Confusion matrix with normalization

Fig. 14 Confusion matrix without normalization

Table 3 Testing data report

Label	Precision	Recall	F1 score	Support
COVID	0.98	0.88	0.93	268
Viral	0.57	0.72	0.64	268
Bacterial	0.72	0.57	0.63	268
Normal	0.83	0.87	0.85	268

Table 4 Comparison between ResNet-18 and ResNet-50

	ResNet-18	Resnet-50
Accuracy (%)	91	76
Sensitivity	0.9	0.7
Specificity	0.96	0.9
Precision	0.9	0.7
F1 score	0.9	0.7

5   Conclusion

In this paper, two models are proposed and compared for detection of COVID-19 along with different classes of pneumonia in order to have more accurate results. The proposed models are ResNet-18 and ResNet-50. The data sets are pre-processed in order to avoid the over fitting and under fitting problems. The ResNet-18 model achieved an accuracy of 91% and ResNet-50 model achieved an accuracy of 75% for same data set. By comparing these two, ResNet-18 model classifies more accurately.

References

1. Das AK, Gosh S, Thunder S, Sachin R, Agarwal automatic COVID-19 detection from X-Ray images using ensemble learning with convolutional neural network
2. Wang L, Lin ZQ, Wong A, COVID—net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images
3. https://github.com/lindawangg/COVID-Net
4. Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A (2020) Covid- caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. arXiv preprint arXiv:2004.02696
5. Ahishali M, Degerli A, Yamac M, Kiranyaz S, Chowdhury ME, Hameed K, Hamid, Mazhar R, Gabbouj M (2020) A comparative study on early detection of covid-19 from chest x-ray images. arXiv preprint arXiv:2006.05332
6. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Y et al (2020) Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest CT. Radiology
7. Jain R, Gupta M, Taneja S, Hemanth DJ, Deep learning-based detection and analysis of COVID-19 on chest X-ray images.

8. Patel P, Chest X-ray (Covid-19& Pneumonia), Accessed at: https:/ww.kaggle.com/prashant268/chest-xray-covid19-pneumonia
9. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China Lancet 395(10223):497–506
10. https://www.kaggle.com/tawsifurrahmancovid19-radiography-database
11. Kaggle datasets download -d Wilhelm Berghammer/chest-xrays-pneumonia-detection