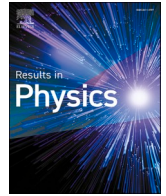




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Automated image classification of chest X-rays of COVID-19 using deep transfer learning

Sara Dilshad^{a,*}, Nikhil Singh^a, M. Atif^b, Atif Hanif^c, Nafeesah Yaqub^b, W.A. Farooq^b, Hijaz Ahmad^d, Yu-ming Chu^{e,f}, Muhammad Tamoor Masood^g

^a Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, Uttar Pradesh, India

^b Department of Physics and Astronomy, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

^c Botany and Microbiology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

^d Section of Mathematics, International Telematic University Uninettuno, Corso Vittorio Emanuele II, 39, Roma 00186, Italy

^e Department of Mathematics, Huzhou University, Huzhou 313000, China

^f Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering Changsha, University of Science & Technology, Changsha 410114, China

^g Institute for Materials Research and Innovation, University of Bolton, Bolton, United Kingdom

ARTICLE INFO

Keywords:

COVID-19 (coronavirus)
Deep learning
Transfer learning
Chest X-Ray (CXR)
Radiology images
Image classification
rRT-PCR
Automated diagnostics

ABSTRACT

Introduction: In December 2019, the city of Wuhan, located in the Hubei province of China became the epicentre of an outbreak of a pandemic called COVID-19 by the World Health Organisation. The detection of this virus by rRTPCR (Real-Time Reverse Transcription-Polymerase Chain Reaction) tests reported high false negative rate. The manifestations of CXR (Chest X-Ray) images contained salient features of the virus. The objective of this paper is to establish the application of an early automated screening model that uses low computational power coupled with raw radiology images to assist the physicians and radiologists in the early detection and isolation of potential positive COVID-19 patients, to stop the rapid spread of the virus in vulnerable countries with limited hospital capacities and low doctor to patient ratio in order to prevent the escalating death rates.

Materials and methods: Our database consists of 447 and 447 CXR images of COVID-19 and Nofindings respectively, a total of 894 CXR images. They were then divided into 4 parts namely training, validation, testing and local/Aligarh dataset. The 4th (local/Aligarh) folder of the dataset was created to retest the diagnostics efficacy of our model on a developing nation such as India (Images from J.N.M.C., Aligarh, Uttar Pradesh, India). We used an Artificial Intelligence technique called CNN (Convolutional Neural Network). The architecture based on CNN used was MobileNet. MobileNet makes it faster than the ordinary convolutional model, while substantially decreasing the computational cost.

Results: The experimental results of our model show an accuracy of 96.33%. The F1-score is 93% and 96% for the 1st testing and 2nd testing (local/Aligarh) datasets (Tables 3.3 and 3.4). The false negative (FN) value, for the validation dataset is 6 (Fig. 3.6), for the testing dataset is 0 (Fig. 3.7) and that for the local/Aligarh dataset is 2. The recall/sensitivity of the classifier is 93% and 96% for the 1st testing and 2nd testing (local/Aligarh) datasets (Tables 3.3 and 3.4). The recall/sensitivity for the detection of specifically COVID-19 (+) for the testing dataset is 88% and for the locally acquired dataset from India is 100%. The False Negative Rate (FNR) is 12% for the testing dataset and 0% for the locally acquired dataset (local/Aligarh). The execution time for the model to predict the input images and classify them is less than 0.1 s.

Discussion and conclusion: The false negative rate is much lower than the standard rRT-PCR tests and even 0% on the locally acquired dataset. This suggests that the established model with end-to-end structure and deep learning technique can be employed to assist radiologists in validating their initial screenings of Chest X-Ray images of COVID-19 in developed and developing nations. Further research is needed to test the model to make it more robust, employ it on multiclass classification and also try sensitise it to identify new strains of COVID-19. This model might help cultivate tele-radiology.

* Corresponding author.

E-mail address: sarabintlateef@gmail.com (S. Dilshad).

<https://doi.org/10.1016/j.rinp.2021.104529>

Received 13 December 2020; Received in revised form 4 July 2021; Accepted 5 July 2021

Available online 28 July 2021

2211-3797/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

COVID-19 has caused a global alarm, broken families, crushed economies, reduced income and burdened the healthcare systems. Coronaviruses are a large family of viruses with some causing less-severe diseases, such as the common cold, and others more severe diseases such as SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus) of 2002 and MERS-CoV (Middle East Respiratory Syndrome Coronavirus) of 2012. In December 2019, the city of Wuhan, located in the Hubei province of China became the epicentre of an outbreak of a pneumonia of an unknown cause [1]. The pneumonia spread quickly, and it was reported at an early stage that patients had contact history with the Huanan seafood market. On 7 January 2020, in a throat swab sample of a patient by the Chinese Centre for Disease Control and Prevention [2], the pathogen of this disease was confirmed by molecular methods as novel coronavirus and WHO temporarily called it as “2019-nCoV acute respiratory disease” [3,4]. On 11th, February 2020 the International Committee on Taxonomy of Viruses (ICTV) announced the official name of the virus to be “SARS-CoV-2” (Sever Acute Respiratory Syndrome Coronavirus-2) due to genetic similarity of the virus to SARS-CoV outbreak of 2002. From a risk of communication perspective and to avoid stigmatisation of regions and ethnic groups, the WHO, on the same date announced the name of this new disease as Coronavirus Disease 2019 “COVID-19” [4,5]. This is the 3rd coronavirus outbreak in the past 20 years and the 6th Public Health Emergency of International Concern (PHEIC) declared by WHO since the International Health Regulations (IHR) came into force [3]. The main transmission routes of COVID-19 identified are respiratory droplets and direct contact with symptomatic and asymptomatic persons. The incubation period was observed around 3–7 days with a maximum of 14 days [6]. With the virus and clinical research moving at breakneck pace, more and more rare and unusual symptoms are coming under the light for different age cohorts that may be associated to SARS-CoV-2 virus. The common symptoms of COVID-19 are cough, shortness of breath or difficulty breathing, fever, chills, muscle pain, sore throat, new loss of taste or smell, other less common symptoms have also been reported, including gastrointestinal symptoms like nausea, vomiting, or diarrhoea [2]. Apart from these, rare and unusual symptoms like multi-system inflammatory syndrome in children, strokes and blood clots in adults, COVID-toes, silent hypoxia and delirium are seen and reported in the Scientist, Exploring Life-Inspiring innovation [7]. The above evidence calls for the need for early diagnosis, isolation and treatment to facilitate research and to flatten the curve by isolating positive patients. Real Time Reverse Transcription-Polymerase Chain Reaction, (rRT-PCR), tests were used for confirming COVID19. However, it is time consuming, have high rates of false negatives between 2% and 29% and the supply of nucleic acid detection kits is also limited [8,9]. The role of imaging in COVID-19 is of paramount importance as the disease’s characteristics manifestations in the lungs show prior to the symptoms [10]. CT scans and X-Rays both can help in early diagnostics and help monitor the clinical course of the disease. CT scans are advanced X-Rays and hence exposes the patient to nearly hundreds of X-Ray radiations. CT scanners need thorough sanitisation after every patient or risk of catching the disease may increase through contamination. CT scanners may not always be readily available to screen the large number of potential COVID-19 patients, especially in developing countries [11]. It is also a challenging task for the radiologists to expertise in the diagnostics of this disease especially in places with limited number of radiologists. It is the need of the hour that the healthcare and artificial intelligence areas merge to prevent unnecessary deaths and to promote tele-radiology by using internet as a primary source to send clinical data and digital images while also following social distancing. Thus, an attempt was made to contribute to this work for early diagnosis of COVID-19 patients amidst the pandemic [12–14]. This study is aimed to establish an early-automated screening model using a low computational transfer learning technique with high accuracy and reduced false negatives by evaluating its performance using

performance metrics, that uses easy to procure and more convenient raw Chest X-Ray (CXR) images to distinguish the COVID-19 cases using deep transfer learning techniques in order to assist the radiologists and help flatten the curve and further test its performance on vulnerable countries like India, with low doctor to patient ratio.

Materials and methods

X-Ray image dataset

The Chest X-Ray (CXR) images in our data set to predict the COVID-19 disease are combined from 2 different sources (Fig. 2.1). The first source is the “covid-chestX-Ray-dataset” which is a public open dataset of CXR and CT images of patients which are positive or suspected of COVID-19 or other viral and bacterial pneumonias (MERS, SARS, and ARDS) developed by Joseph Paul Cohen and is available on GitHub repository. This dataset is compiled from public sources as well as through indirect collection from hospitals and physicians. This dataset is constantly updated and images for the chest X-Ray are in dcm, jpg, or png formats [15]. The second source is the “CXr8” database developed by Ronald M. Summers from National Institutes of Health- Clinical Centre from where we have gathered the images of normal patients/No-findings. These images are available in png format [16]. We have an additional dataset that has been locally and directly obtained from Jawaharlal Nehru Medical College, Aligarh Muslim University (JNMC, AMU), Aligarh, Uttar Pradesh, India, upon which the model shall be tested and the performance metrics analysed. We received these images via e-mail and the details of the patient were cropped out [17]. The first source contained 758 Chest X-Ray (CXR) and CT images of various lung diseases like COVID-19, SARS, and Legionella etc. Total COVID-19 images were 521 from which we dug 417 CXR images with

PA, AP and SupineAP view only. From the second dataset we took 447 normal images or No-finding CXR images. These 417 and 447 images were combined to form our database of 864 CXR images. There was an additional dataset acquired locally from J.N.M.C., A.M.U. that had 30 COVID-19 images, extending our database to 447 COVID-19 and 447 No-finding CXR images, a total of 894 images.

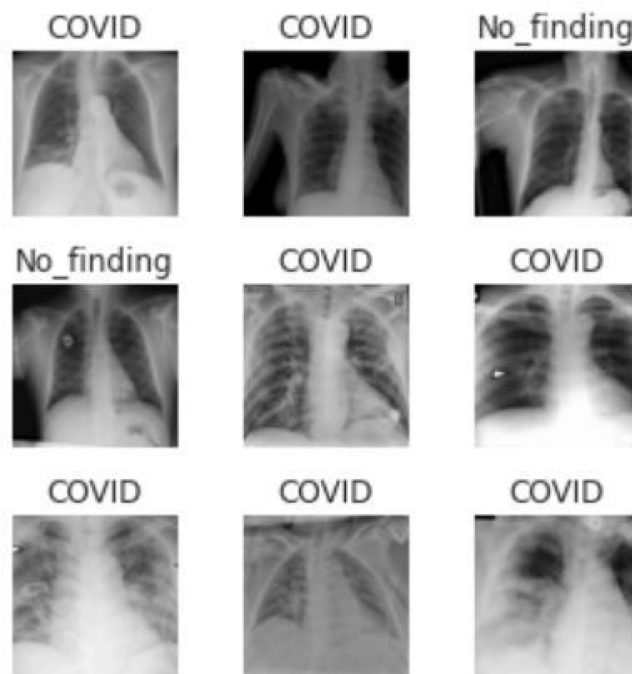


Fig. 2.1. Some COVID and NO-finding CXR images from our database.

Data processing

The database was divided into 3 parts namely training, validation and testing dataset. The 1st (training) folder had 309 COVID-19 images and 320 No-finding images. The 2nd (validation) folder had 58 COVID19 images and 51 No-finding images and the 3rd (testing) folder had 50 COVID-19 images and 49 Nofinding images. The 4th (local/ Aligarh) folder of the dataset was created to re-test the diagnostics efficacy of our model and this folder had 30 COVID-19 images and 27 No-finding images.

The CNN Architecture-MobileNet

Artificial Intelligence has bridged the gap between the capabilities of humans and machines. Computer vision is a domain of AI that enables machines to perceive the world like humans. The advancements in these fields has been done over one particular algorithm called a Convolutional Neural Network. A CNN consists of an input layer, hidden layer(s) and output layer. The hidden layers consist of convolutional layers, feature extraction carried out by additional layers such as activations function, pooling layer, batch normalization and the fully connected layer completes the classification process [18].

To develop a diagnostic model for classifying COVID-19 and non COVID-19 patients, we choose a model based on deep learning MobileNet Convolutional Neural Network. MobileNet is considered much faster than standard convolutional network due to its distinct filter approach to each response channel. Our model is constructed on depth wise separable convolution which has two succeeded functions, one is a depth wise convolution at filtering stage: that applies convolution to single input channel at a time, second is a point wise convolution at filtering stage: that performs linear combination of outputs to the depth wise convolution (Fig. 2.2).

Batch normalization and rectified linear unit (ReLU) layer come after each convolution stage. Computational cost get reduces phenomenally in depthwise separable approach due to separate filtration at combining steps to minimize the size of the model and its complexity. An example is as explained:

$$Df \times Df \times M$$

For a feature map of $Df \times Df$ in size, the kernel size is $Dk \times Dk$, the input channel is M , the output channel is N . The computational cost of the standard convolutions can be seen in this equation:

$$Dk \times Dk \times M \times N \times D \times Df,$$

The theory above simplifies MobileNet and makes it faster than the ordinary convolutional model, and thus, decreases the computational cost [19].

The Version used for this model was MobileNet_V2 with 3.47 million parameters. The multiplyaccumulate operations (MACs) were 300 million.

Results

The evaluation platform used was the Kaggle cloud service platform with CUDA version 9.2 and the evaluation time taken was 20 min using the 17.1 GB NVIDIA Tesla P100 GPU.

Fig. 3.1 shows the boxplot of the ages of COVID-19 patients. Fig. 3.2 shows the boxplot of the ages of No-finding patients. The boxplot indicates nearly identical ages of patients from both categories. The summary of the two plots of COVID-19 and No-finding patients are shown in Table 3.1 and Table 3.2 respectively. The range of COVID19 lies from 12 to 94. The range of No-findings patient images lies from 20 to 80. The mean of COVID-19 and No-findings patient images are 58.73 and 48.48 respectively. The median of COVID-19 and No-findings patient images are 60 and 50 respectively. The Q1 and Q2 values of COVID-19 are 46 and 72. The Q1 and Q2 values of No-finding patient images are

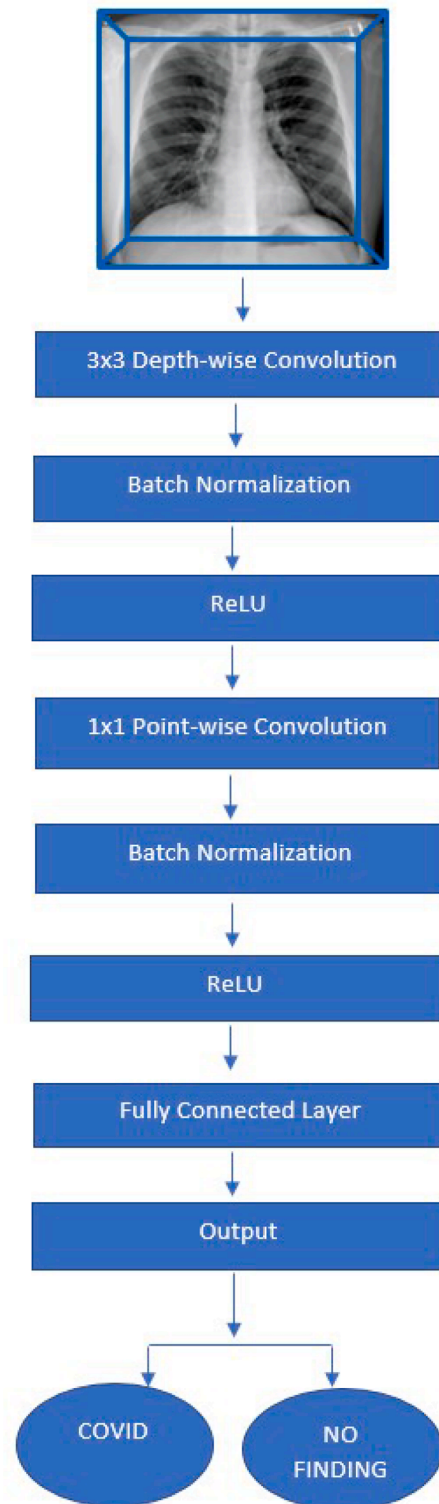


Fig. 2.2. A flow chart representing the CNN Architecture, MobileNet.

37 and 59 respectively. Confusion Matrix is a performance measurement for machine learning where output can be of two or more classes. It gives you an insight not only into the errors being made by your classifier but more importantly the types of errors that are being made. A true positive (TP) is an outcome where the model correctly predicts the positive class. Similarly, a true negative (TN) is an outcome where the model correctly predicts the negative class. A false positive (FP) is an outcome where the

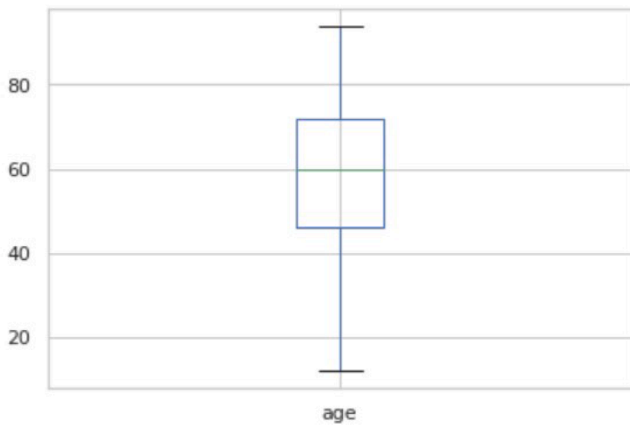


Fig. 3.1. Boxplot of the ages of COVID-19 patients.

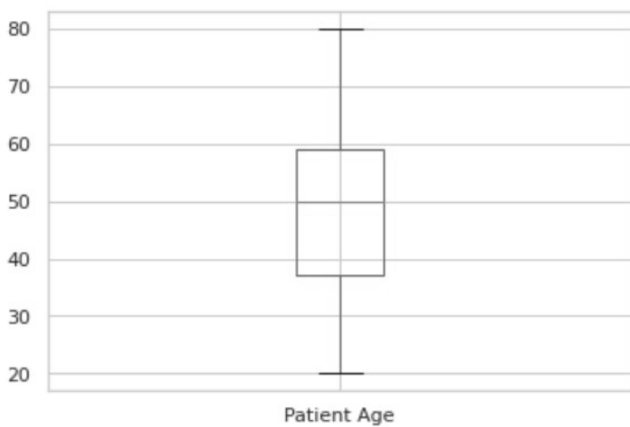


Fig. 3.2. Boxplot of the ages of No-finding patients.

Table 3.1
Summary of the COVID-19 patients.

Count	185.000000
Mean	58.729730
Std	16.312586
Min	12.000000
25%	46.000000
50%	60.000000
75%	72.000000
Max	94.000000

Table 3.2
Summary of the No-finding patients.

Count	447.000000
Mean	48.483221
Std	14.215796
Min	20.000000
25%	37.000000
50%	50.000000
75%	59.000000
Max	80.000000

model incorrectly predicts the positive class, also called as type 1 error. And a false negative (FN) is an outcome where the model incorrectly predicts the negative class, also called as type 2 error.

Accuracy: The total number of true predictions in total dataset. It is represented by the equation of true positive and true negative examples divided by true positive, false positive, true negative and false negative.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: Precision is interested in the number of genuinely positive examples your model identified against all the examples it labelled positive. Mathematically, it is the number of true positives divided by the true positives plus the false positives.

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

Recall/Sensitivity: The term recall, sensitivity or true positive rate refers to the proportion of genuine positive examples that a predictive model has identified. To put that another way, it is the number of true positive examples divided by the total number of positive examples and false negatives.

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

F1-score: F-score helps to measure Recall and Precision at the same time. It uses Harmonic Mean in place of Arithmetic Mean by punishing the extreme values more.

$$f1 \text{ score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

AUC-ROC Curve: An AUC-ROC curve is a performance measurement for classification problem at various thresholds settings. It is also called as c-statistic. ROC is a probability curve and AUC represent degree or measure of separability. It tells how much model is capable of distinguishing between classes. Higher the AUC, better the model is at predicting 0s as 0s and 1s as 1s. By analogy, Higher the AUC, better the model is at distinguishing between patients with disease and no disease. The ROC curve is plotted with TPR against the FPR where TPR is on y-axis and FPR is on the x-axis.

- True Positive Rate (TPR): It is a synonym for recall and is therefore defined as

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

- False Positive Rate (FPR): It is defined as

$$FPR = 1 - Specificity$$

$$FPR = \frac{FP}{FP + TN}$$

- False Negative Rate (FNR): It is defined as

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

The ROC area for validation, testing and local/Aligarh dataset are 0.9936, 0.9298 and 0.9630 as shown in Figs. 3.3, 3.4 and 3.5, respectively. From Fig. 3.6 we know that for validation dataset TP = 56, FN = 2, FP = 2, and TN = 49. From Fig. 3.7 we see that for testing dataset TP = 44, FN = 6, FP = 1 and TN = 48 and from Fig. 3.8 we know that for local/Aligarh dataset TP = 30, FN = 0, FP = 2 and TN = 25. The classification report in Table 3.3 for the testing dataset shows an F1-score of 93%. The recall/sensitivity for detection of COVID-19 specifically is 88%. The macro-average and weighted-average recall of the classifier are both 93%. The classification report in Table 3.4 for the local/Aligarh dataset manifests an F1-score of 96%. The recall/sensitivity for detection of COVID-19 specifically is 100%. The macro-average and weighted-average recall of the classifier are both 96%. The False Negative Rate (FNR) calculated from the above formula is 12% for the testing dataset and 0% for the locally acquired testing dataset (local/Aligarh).

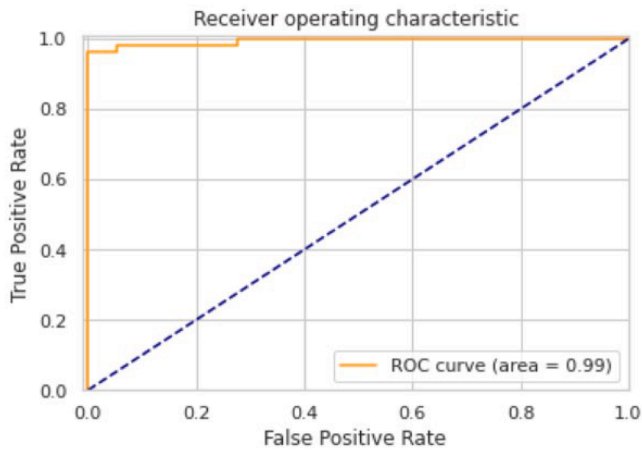


Fig. 3.3. The AUC-ROC curve of the validation dataset.

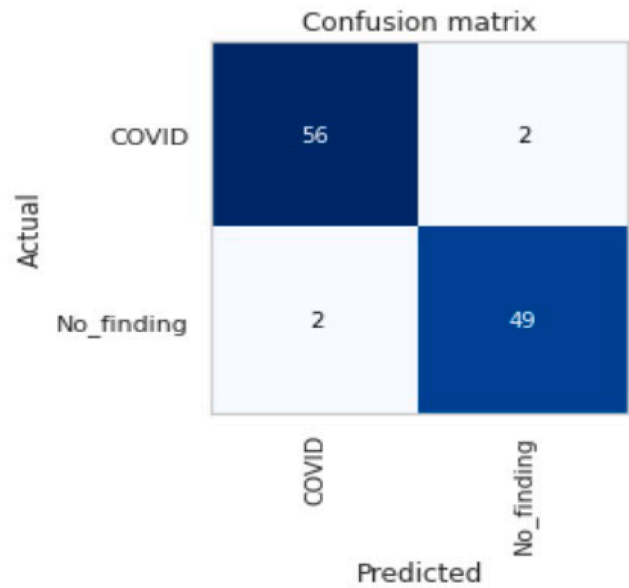


Fig.3.6. The co confusion matrix of the validation dataset.

The co confusion matrix of the validation dataset.

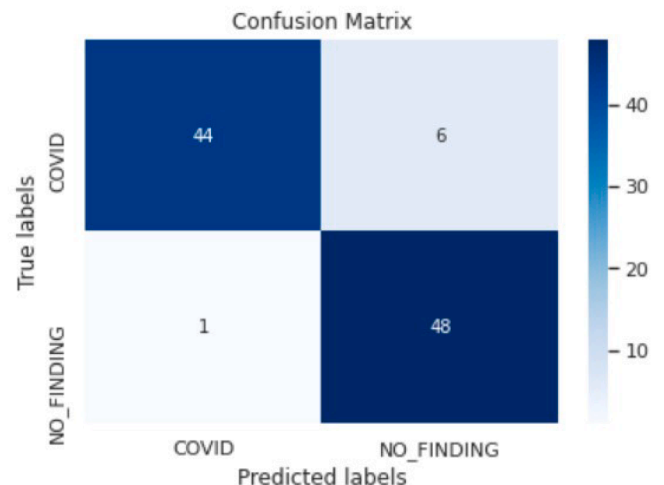


Fig. 3.7. The confusion matrix of the testing dataset.

$$FNR(testing) = \frac{FN}{FN + TP} = \frac{6}{6 + 44} = \frac{6}{50} = 0.12$$

$$FNR(local/Aligarh) = \frac{FN}{FN + TP} = \frac{0}{0 + 30} = \frac{0}{30} = 0.00$$

Accuracy is considered important when true positives and true negatives are key. It is a useful measurement of performance when the classes are balanced. In the case where false positives and false negatives are of immense significance, F1-score is a better metric. This is also useful when the classes are imbalanced. We will tend to focus slightly more on the F1-score than accuracy, although in our case both are equally important.

From Table 3.5 we see that the model has an error rate of 3.67% for the 20th epoch. From Table 3.6 we see that the top 1 accuracy of our model is 96.33%. This means that the first guess of our model is 96.33% accurate and the model gives 100% accurate prediction results on the second guess. The optimal learning rate for our model lies between $1e^{-03}$ and $1e^{-02}$ as shown in Fig. 3.9, where it gives the least loss. Learning rate

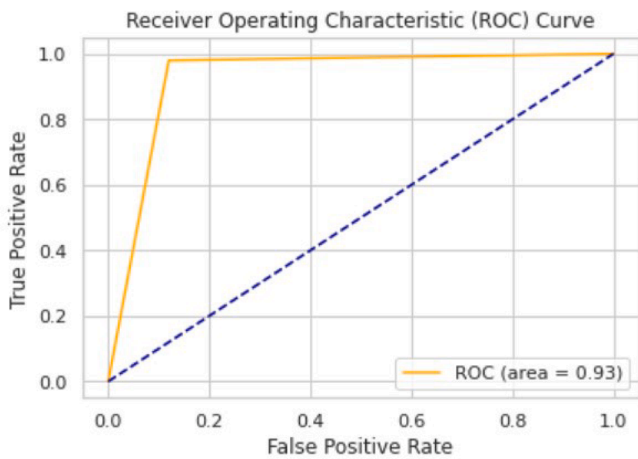


Fig. 3.4. The AUC-ROC curve of the testing dataset.

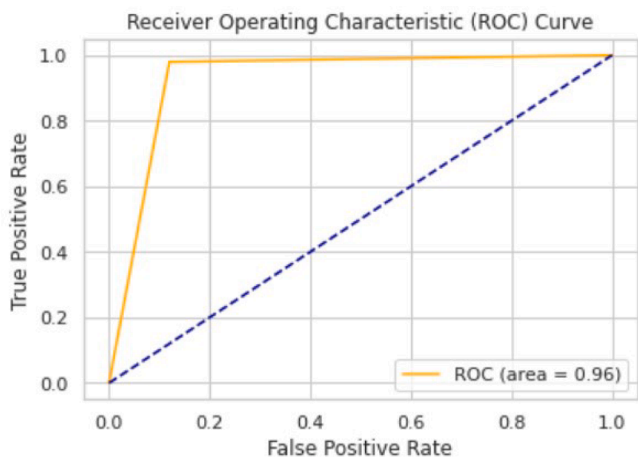


Fig. 3.5. The AUC-ROC curve of the local/Aligarh dataset.

The confusion matrix of the testing dataset.

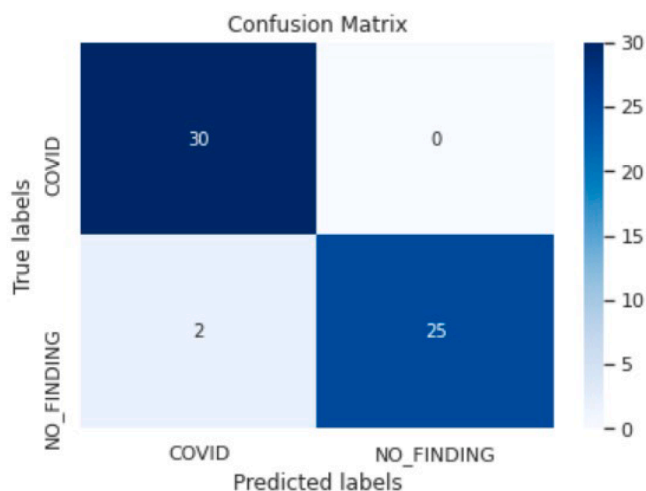


Fig. 3.8. The confusion matrix of the local/Aligarh dataset.

Table 3.3

The classification report of the testing dataset.

	Precision	Recall	F1-Score	Support
COVID	0.98	0.88	0.93	50
No Finding	0.89	0.98	0.93	49
Accuracy			0.93	99
Macro Avg	0.93	0.93	0.93	99
Weighted Avg	0.93	0.93	0.93	99

Table 3.4

The classification report of the local/Aligarh dataset.

	Precision	Recall	F1-Score	Support
COVID	0.94	1.00	0.97	30
No Finding	1.00	0.93	0.96	27
Accuracy			0.96	57
Macro Avg	0.97	0.96	0.96	57
Weighted Avg	0.97	0.96	0.96	57

Table 3.5

Table of the error rate of the model.

Epoch	Train_Loss	Valid_Loss	Error_Rate	Time
0	0.561701	1.334554	0.339450	00:39
1	0.415436	0.222622	0.091743	00:39
2	0.320269	0.174101	0.045872	00:38
3	0.259764	0.400186	0.082569	00:39
4	0.210110	1.650870	0.155963	00:37
5	0.201729	0.709523	0.082569	00:39
6	0.199924	0.671769	0.119266	00:40
7	0.182643	0.425020	0.082569	00:40
8	0.178835	0.290745	0.064220	00:39
9	0.160731	0.079751	0.055046	00:38
10	0.145512	0.292171	0.045872	00:38
11	0.146422	0.109115	0.027523	00:36
12	0.144602	0.154809	0.036697	00:39
13	0.160723	0.597565	0.128440	00:38
14	0.160858	0.178486	0.064220	00:37
15	0.153326	0.294835	0.073394	00:38
16	0.152694	0.125383	0.055046	00:37
17	0.156151	0.123060	0.036697	00:38
18	0.141335	0.172965	0.045872	00:38
19	0.126174	0.140441	0.036697	00:38

Table 3.6

Table of accuracy of the validation dataset.

Top 1 Accuracy	0.963302731513977
Top 2 Accuracy	1.0

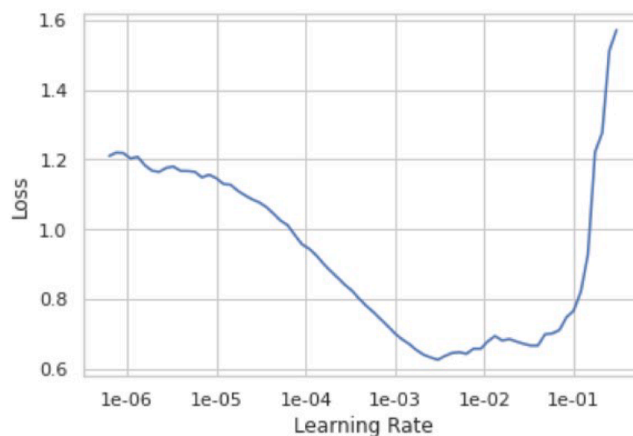


Fig. 3.9. Graph depicting the learning rate of the model.

Graph depicting the learning rate of the model.

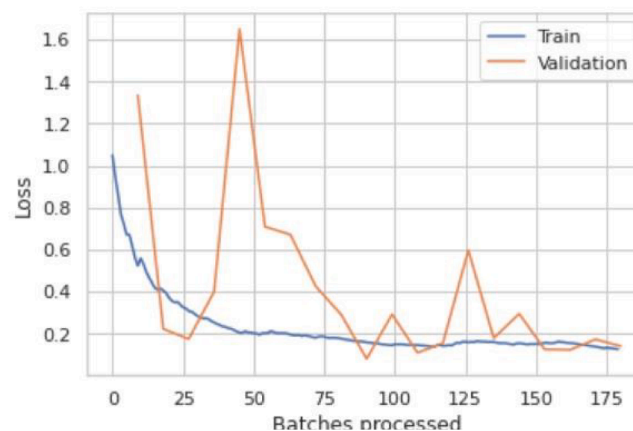


Fig. 3.10. The loss function graph of the training and validation dataset.

is a hyper-parameter while configuring your neural network that controls how much to change the model in response to the estimated error each time the model weights are updated. Small values can result in longer training process of the model whereas higher values could lead to an unstable training process. Fig. 3.10 shows the loss function for training and validation dataset. The loss functions tracks loss during training over time as it is evaluated on the individual batches during the forward pass. For most of our training we obtained a good learning rate with a low value of loss. During validation we see a couple of peaks but with a continuous decay in loss which eventually falls down to a small low value of loss. Thus, the results for loss function was a good “flattened loss”.

Figs. 3.11, 3.12, 3.13, 3.14, 3.15 and 3.16 all show the predicted output class for each of the input raw CXR images. Figs. 3.14, 3.15 and 3.16 are the predicted class of the locally acquired images from Aligarh, U.P., India. The execution time for the model to predict the input images and classify them is less than 0.1 s. The time taken by the classifier to classify two random images was 0.08546829223632812 s and 0.08694672584533691 s.

Our model compared to other binary classification models using

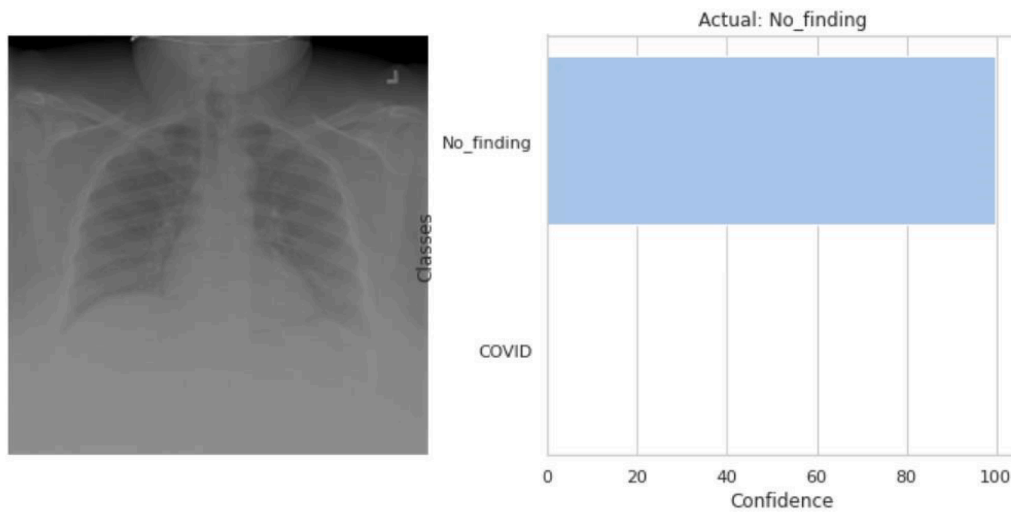


Fig. 3.11. The model predicted that the input CXR image (left) belongs to the No_finding's class as shown in the output (right).

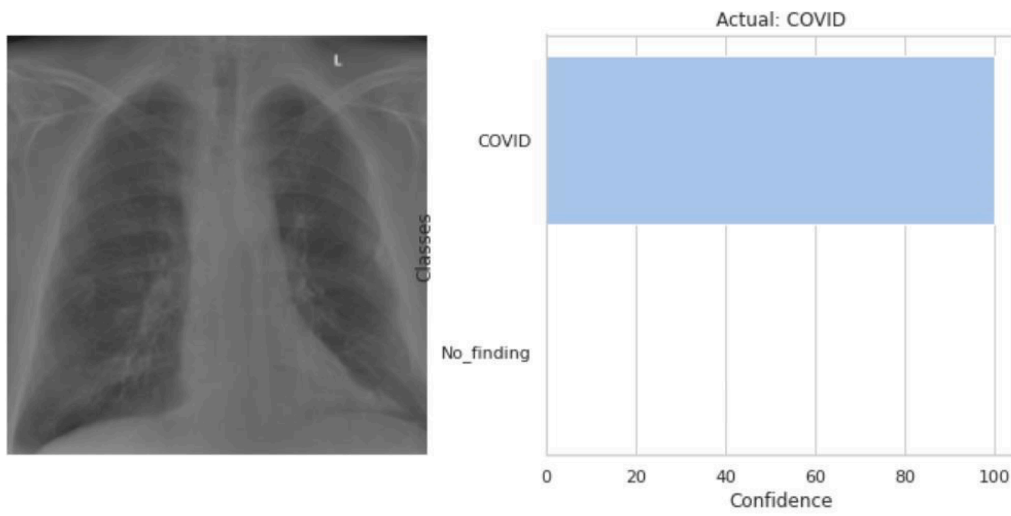


Fig. 3.12. The model predicted that the input CXR image (left) belongs to the COVID class as shown in the output (right).

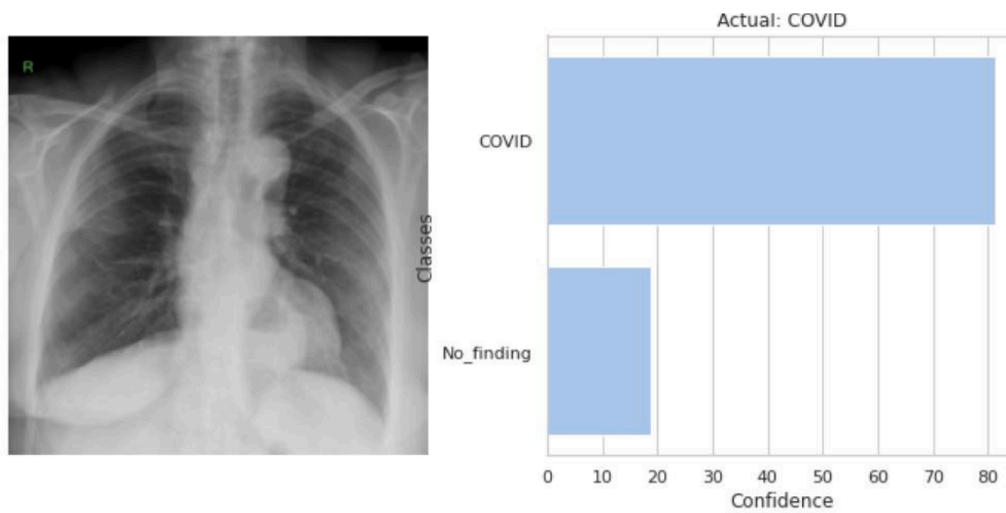


Fig. 3.13. The model predicted that the input CXR image (left) belongs to the COVID class as shown in the output (right).

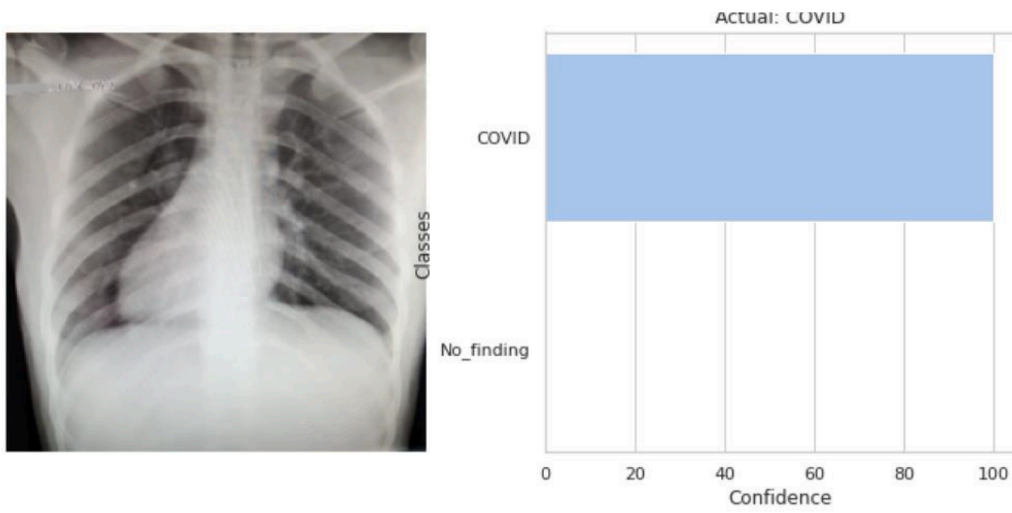


Fig. 3.14. The model predicted that the input CXR image (left) belongs to the COVID class as shown in the output (right).

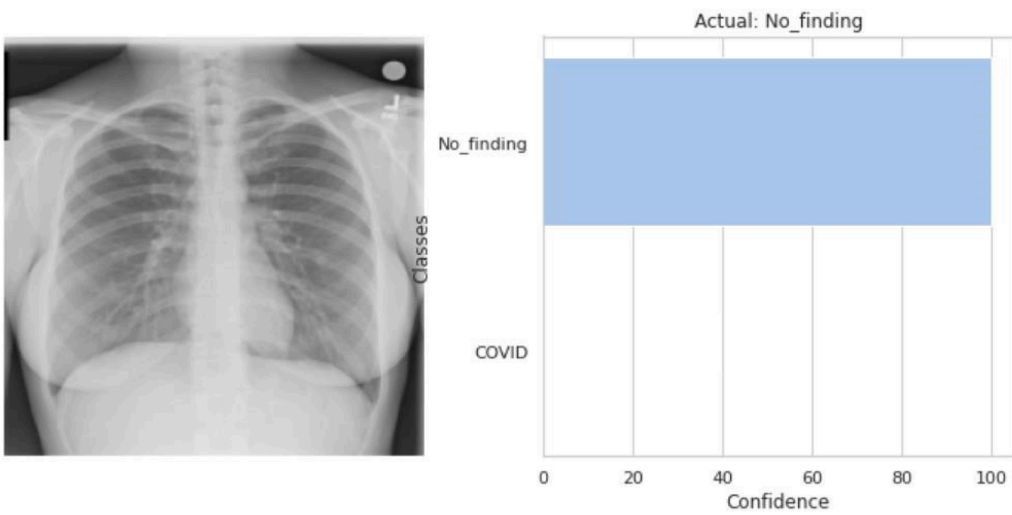


Fig. 3.15. The model predicted that the input CXR image (left) belongs to the NO_finding's class as shown in the output (right).

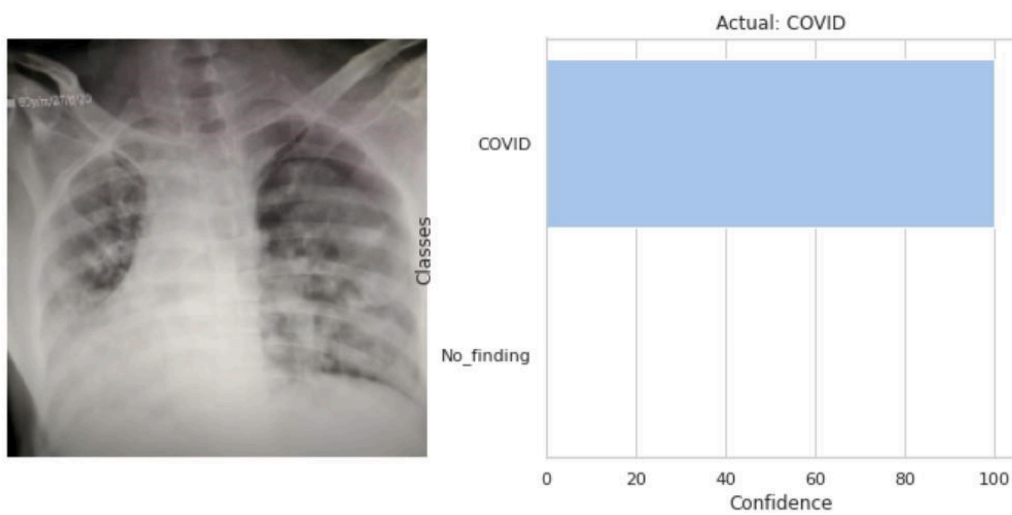


Fig. 3.16. The model predicted that the input CXR image (left) belongs to the COVID class as shown in the output (right).

radiological imaging for the diagnosis of COVID-19 is shown below.

Discussion

The novel coronavirus that emerged in Wuhan has now submerged the entire globe. It has exposed the broken healthcare systems and lack of healthcare facilities. The gold standard nucleic acid-based detection has its shortcomings including high rate of false negatives due to factors like methodological disadvantages, disease stages etc [20]. Many open sources of COVID-19 images have emerged for the public to use for the greater good. Ying S. et al. (2020) proposed DRE-Net model built on the pre-trained ResNet50 and attained an accuracy of 86% using CT scans [12]. They used 777 COVID-19 images and 708 Healthy images. Wang et al. (2020) achieved an accuracy of 82.9% using CT scans and a deep learning model called modified Inception (M-Inception), using 195 COVID-19 positive images and 258 negative images [20]. Hemden et al. (2020) proposed COVIDX-Net model to identify positive patients using X-Rays to achieve an accuracy of 90%, with 25 COVID-19 images and 25 normal images [21]. (Sethy and Behera 2020) used X-Rays for their model ResNet + SVM and got an accuracy 95.38%, from 25 COVID-19 positive images and 25 negative images [22]. Narin et al. (2020) used 50 COVID-19 positive and 50 negative images and achieved an accuracy of 98% for their proposed model [23]. Using UNet + 3D Deep Network model Zheng et al. (2020) achieved an accuracy of 90.80% using images of chest CT from 313 COVID-19 positive and 229 negative COVID-19 images [24]. A.S. Al-Waisey et al. (2020) described another hybrid model COVID-CheXNet that reached an accuracy of 99.99% using 400 confirmed COVID-19 infection images and 400 normal condition images [25]. These studies comparable to our model can be viewed in Table 3.7. Our data is taken from two open sources and it consists of 864 CXR images. Our study uses Chest X-Ray (CXR) type images with 417 COVID-19 and 447 No-finding CXR images. Our utmost concern was to have CXR images of comparable age groups of people as age can have a significant impact on the diagnostic ability of the automated model. Tulin Ozturk et al. (2020), Shumaila Javeed et al. (2021) attained an accuracy of 98.08% for a three outcome classifier (COVID, No-Findings and Pneumonia) [26,27]. This accuracy could be attributed to use of non-COVID images of children dataset belonging to the age 1–5 years, due to the lack of availability or restriction on adult image datasets during the time of their study. Children CXR images are evidently different from adult CXR images which if used can cause the accuracy to shoot up due to obvious differentiation between the two. Hence, taking data of similar age cohorts is indispensable. Our proposed model attained a high accuracy of 96.33%. Since our data was balanced the accuracy does play an important role in the justification of our model. As we have the data of an infectious disease, it is required that we reduce the number of false negatives i.e. positive COVID19 patients being predicted as negative. The FN value of our model was exceptionally low. Recall is also a useful metric in cases where false negative trumps false positive as we accidentally do not want to discharge an infectious person and have them mixed with the healthy population thereby spreading the contagious virus. The value of F1-score which is the harmonic mean of precision and

Table 3.7

Comparison of the proposed model with various other binary classification models.

Study	Image Type	Method Used	Accuracy
<i>Proposed Model</i>	Chest X-Ray	MobileNet_V2	96.33%
Ying et al.	Chest CT	DRE-Net	86.00%
Wang et al.	Chest CT	M-Inception	82.90%
Hemdan et al.	Chest X-Ray	COVIDX-Net	90.00%
Sethy and Behra	Chest X-Ray	ResNet50 + SVM	95.38%
Narin et al.	Chest X-Ray	Deep CNN ResNet-50	98.00%
Zheng et al.	Chest CT	UNet + 3D Deep Network	90.80%
A.S. Al-Waisy et al.	Chest X-Ray	COVID-CheXNet	99.99%

recall is 93% for the testing dataset and 96% for the second testing (local/Aligarh) dataset. For the 1st testing dataset, we obtained the recall value of the classifier as 93% and a recall value for the class of all COVID-19 images as 88%. For the 2nd (local/Aligarh) testing dataset, we obtained the recall value of the classifier as 96% and a recall value for the class of all COVID-19 images as 100%. The False Negative Rate (FNR) is 12% for the testing dataset and 0% for the locally acquired dataset (local/Aligarh). According to another study, for a 100% assumed specificity of the diagnostic assay the FNR was 9.3% and sensitivity/recall was 90.7% and it was suggested that rRT-PCR results alone should not be the deciding factor for COVID-19 [28]. This means that the false negative rate is quite low for the above mentioned datasets.

Henceforth, all the above performance metrics indicates that with a few more training and tests this model can be employed in hospitals as aimed by the paper.

The advantages of using this technique model are:

- 1) Higher accuracy
- 2) Extremely low false negatives
- 3) Classification of raw images of data
- 4) Fully automated end-to-end (back-end-to-front-end) structure
- 5) No feature extraction required
- 6) Results in less than a second
- 7) X-Ray images are easier to procure than CT scans, especially in developing countries.
- 8) Patient is exposed to lesser radiation in X-Ray
- 9) X-Ray units are easier to sanitise than CT scanners
- 10) An approach to assist radiologists
- 11) Promote tele-radiology while abiding the social distancing protocol

Due to the COVID-19 outbreak and rising cases in the world, especially India, our study was limited from being clinically tested. We could collect only 30 local radiology images for COVID-19 cases and evaluate them with our model. In future, we aim to validate the model by incorporating more CXR images and will further take the model a step ahead from binary classification to multi-class classification. We might even experiment with different layering structures and compare the results. After conducting clinical trials, we aim to deploy the model in the local hospitals for early screening of potential COVID-19 patients. We aim to further test the model on multiple classification and increase its sensitivity on the different variants of COVID-19.

Conclusion

In this study, we established an early screening and fully automated model with end-to-end structure without necessitating feature extraction for the detection of COVID-19 from Chest X-Ray images with deep learning based technologies and an accuracy of 96.33%. The value of recall/sensitivity for testing and local/Aligarh dataset was 93% and 96% respectively. The False Negative Rate (FNR) is 12% for the testing dataset and 0% for the locally acquired dataset (local/Aligarh). Imaging of COVID-19 with X-Rays is more feasible in developing countries and wherever the patient count surpasses existing imaging modalities. This automated model can help reduce patient load for radiologists. It could be a promising supplementary aid for frontline workers typically in countries where more and more healthcare workers have been isolated after having tested positive causing an acute shortage of health care workers. Social distancing norms are also fulfilled as this technology promotes tele-radiology. COVID-19 has burdened the healthcare systems and economies. Early diagnosis with the aid of image classification models allows early containment of this contagious disease and assists in flattening the curve. We intend to make our model more robust and accurate by validating it with additional images database.

Clinical trial

No clinical trial conducted. Images from Jawaharlal Nehru Medical College, Aligarh Muslim University (JNMC, AMU), Aligarh, and Uttar Pradesh, India were tested to measure the diagnostic efficacy of the model.

Informed consent and patient details

Images were mostly taken from online open data sources. The local/ Aligarh images were taken without knowing the details of the patient. Further possible traces of the patient on the CXR images were also removed before usage.

CRediT authorship contribution statement

Sara Dilshad: Conceptualization, Visualization, Writing - original draft, Data curation, Formal analysis. **Nikhil Singh:** Conceptualization, Visualization, Writing - original draft, Data curation, Formal analysis. **M. Atif:** Writing - review & editing, Formal analysis, Funding acquisition. **Atif Hanif:** Writing - review & editing, Formal analysis, Funding acquisition. **Nafeesah Yaqub:** Writing - review & editing, Formal analysis, Funding acquisition. **W.A. Farooq:** Writing - review & editing, Formal analysis, Funding acquisition. **Hijaz Ahmad:** Writing - review & editing, Formal analysis, Funding acquisition. **Yu-ming Chu:** Writing - review & editing, Formal analysis, Funding acquisition. **Muhammad Tamoor Masood:** Writing - review & editing, Formal analysis, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Researchers Supporting Project number (RSP-2021/397), King Saud University, Riyadh, Saudi Arabia.

References

- [1] WHO, statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china, (2020). <https://www.who.int/china/news/detail/09-01-2020>.
- [2] CDC, Centre for disease and control and prevention, (2021). <https://www.cdc.gov/coronavirus/2019ncov/symptoms-testing/symptoms.html>.
- [3] Ge H, Wang X, Yuan X, Xiao G, Wang C, Deng T, et al. The epidemiology and clinical information about COVID-19. *Eur J Clin Microbiol Infect Dis* 2020;39(6): 1011–9. <https://doi.org/10.1007/s10096-020-03874-z>.
- [4] WHO, world health organization, Novel Coronavirus (2019-nCoV) Situation Report-10; 2020.
- [5] WHO, World Health Organization, naming the coronavirus disease as covid -19, (2020). <http://www.euro.who.int/en/health-topics/health-emergencies/international-healthregulations/news/news/2020/2/2019-ncov-outbreak-is-an-emergency-of-international-concern>.
- [6] Odhiambo JO, Ngare P, Weke P, Otieno RO. Modelling of COVID-19 Transmission in Kenya Using Compound Poisson Regression Model. *J Adv Math Comput Sci* 2020;101–11. <https://doi.org/10.9734/jamcs/2020/v35i230252>.
- [7] The Unusual symptoms of Covid 19. Exploring Life, Inspiring Innovation: The Scientist, (2020). <https://www.the-scientist.com/news-opinion/the-unusual-symptoms-of-covid-19-67522>.
- [8] Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, et al. False-negative results of initial RTPCR assays for Covid-19 a systematic review. *PLoS ONE* 2020;15(12):e0242958. <https://doi.org/10.1371/journal.pone.0242958>.
- [9] Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol* 2020;126:108961. <https://doi.org/10.1016/j.ejrad.2020.108961>.
- [10] Yang W, Sirajuddin A, Zhang X, Liu G, Teng Z, et al. The role of imaging in 2019 novel coronavirus Pneumonia (COVID-19). *Eur Radiol* 2020;30(9):4874–82. <https://doi.org/10.1007/s00330-020-06827-4>.
- [11] Jacobi A, Chung M, Berheim A, Eber C. Portable chest X ray in Coronavirus diseases-19 (COVID-19): a pictorial review. *Clin Imag* 2020;64:35–42. <https://doi.org/10.1016/j.clinimag.2020.04.001>.
- [12] Song Y, Zheng S, Li L, Zhang X, Zhang X, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *medRxiv* 2020. <https://doi.org/10.1101/2020.02.23.20026930>.
- [13] [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it).
- [14] WHO, Director-General Opening Remarks at media briefing on COVID-19, (2020). <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19>.
- [15] <https://github.com/ieee8023/covid-chestX-Ray-dataset>.
- [16] <https://nihcc.app.box.com/v/ChestX-Ray-NIHCC/folder/36938765345>.
- [17] <https://github.com/singhn156/Covid-Aligarh>.
- [18] Mohammed MA, Abdulkareem KH, Garcia-Zapirain B, Mostafa SA, Maashi MS et al. A comprehensive investigation of machine learning feature extraction and classification methods for automated diagnosis of COVID-19 based on X-ray images. 66(3) (2021), 3289-3310. doi:<https://techscience.com/cm/v66n3/41053>.
- [19] Bi C, Wang J, Duan Y, Fu B, Kang J-R, Shi Y. MobileNet based apple leaf diseases identification. *Mobile Netw Appl* 2020. <https://doi.org/10.1007/s11036-020-01640-1>.
- [20] Wang S, Kang B, Ma J, Zeng X, Xiao M, et al. A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). *Eur Radiol* 2021;31(8): 6096–104. <https://doi.org/10.1007/s00330-021-07715-1>.
- [21] Hemdan EED, Shouman MA, Karar ME. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images, (2020), *ArXiv*, abs/2003.11055.
- [22] Sethy PK, Behera SK. Detection of Coronavirus disease (COVID-19) based on deep features. *Preprints* 2020. <https://doi.org/10.20944/preprints202003.0300.v1>. 2020030300.
- [23] Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using XRay images and deep convolutional neural networks. *Pattern Anal Appl* 2021. <https://doi.org/10.1007/s10044-021-00984-y>.
- [24] Zheng C, Deng X, Fu Q, Zhou Q, Feng J, et al. Deep learning-based detection for COVID-19 from chest CT using weak label. *medRxiv*, (2020). doi: <https://doi.org/10.1101/2020.03.12.20027185>.
- [25] Al-Waisy AS, Al-Fahdawi S, Mohammed MA, Abdulkareem KH, Mostafa SA, et al. COVID-CheXNet: hybrid deep learning framework for identifying COVID-19 virus in chest X-rays images. *Soft Comput* 2020:1–16. <https://doi.org/10.1007/s00500-020-05424-3>.
- [26] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 2020;121:103792. <https://doi.org/10.1016/j.combiomed.2020.103792>.
- [27] Shumaila J, Anjum S, Alimgeer KS, Atif M, Khan S, et al. A novel mathematical model for COVID-19 with remedial strategies. *Results Phys* 2021;27:104248. <https://doi.org/10.1016/j.rinp.2021.104248>.
- [28] Kanji JN, Zelyas N, MacDonald C, Pabbaraju K, Khan MN, et al. False negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virology J* 2021;18(1). <https://doi.org/10.1186/s12985-021-01489-0>.