COVID-19 Detection System From Chest X-Rays Using Deep Learning & Integrating Firebase Database

Majid Khan[1], Hayyan Ansari[2], Abhyudaya Tiwari[3], Chetan Amrutkar[4], Prof. Vishakha Bhadane[5]
[1, 2, 3, 4, 5] Jawahar Education Society's, Institute of Technology, Management & Research, Nashik,

Abstract- *The outstanding expansion in COVID-19 patients is overpowering medical care frameworks across the world. With restricted testing units, it is incomprehensible for each understanding with respiratory ailment to be tried using current regular methods (RT-PCR). The tests also have a long turn-around time as they take hours to days to fetch results. Recognizing conceivable COVID-19 contamination on Chest X-Ray might help isolate high-hazard patients while test results are anticipated. X-Ray machines are now accessible in most medical care frameworks, and with most present-day X-Ray frameworks previously digitized, there is no transportation time required for the examples by the same token. In this work, we propose the utilization of chest X-rays to focus on the determination of patients for additional RT-PCR testing. This might be helpful where frameworks are attempting to conclude whether to keep the patient in the ward alongside different patients or disengage them in COVID-19 regions. It would also help in identifying patients with a high likelihood of COVID-19 with a false negative RT-PCR who would need retesting. Further, we propose the utilization of current AI strategies to distinguish the COVID-19 patients utilizing X-Ray images in an automated way, especially in settings where radiologists are not available, and help make the proposed testing technology scalable. We present COVID-19 Detector, a deep learning model integrated into a web application for precise testing. On the publicly available COVID-19 chest X-Ray dataset provided by World Health Organization (WHO), our model gives about 99% accuracy for the COVID-19 infection.*

Keywords- COVID-19, RT-PCR, Chest X-rays, Deep learning, Web application.

I. INTRODUCTION

COVID-19 (Coronavirus Disease 2019) is an infectious illness caused by a large family of bat viruses that had never been found previously in humans. This has led to a worldwide pandemic since February 2020. Coronavirus (CoV) is a group of many infections that typically contaminate creatures like chickens, bats, camels, felines. These infections inside the creatures can transform allowing them to send the infection to different species prompting an overflow. It has been seen that the infection is assaulting the respiratory framework in people, changing from the normal virus to all the more destructive illnesses like Middle East Respiratory Syndrome and Severe Acute Respiratory Syndrome (SARS). The side effects of COVID-19 can be smooth; a few groups may not show any manifestations whatsoever however can taint others. The ordinary indications of COVID-19 are fever, dry cough, and weakness. Additional symptoms that are less common in nature are headaches, nasal congestion, sore throat, loss of taste and smell, conjunctivitis, and discoloration of fingers or toes. Typically, more seasoned individuals having other hidden clinical issues like hypertension, heart and lung issues, diabetes, or malignancy are at a higher danger of fostering the extreme disease. However, anyone can get the virus and get ill. In many developed countries, the health system has come to the point of collapse due to the accelerating demand for ICUs (Intensive Care Units) simultaneously. Concentrated consideration units are loaded up with patients who deteriorate with COVID-19.

The virus is advancing its spread across the world, with more than 18Cr confirmed cases in 188 countries, as of June 27th, 2021. Europe and North America saw the first significant outbreaks in April, but as they drove to ease, Latin America and Asia started noticing an escalation in cases. North America has seen a resurgence of diseases as of late, generally coordinated by new episodes in the US, however, Mexico has likewise seen a rising number of cases. Not just this, COVID-19 has gotten a worldwide stop monetary movement. The economic damage is evident. This pandemic is one of the biggest financial stuns the world has encountered in many years. Late information shows that the help business has been hit the hardest. The travel industry and worldwide exchange have been greatly upset in each locale of the world.

There is a desperate need for an automated testing system for COVID-19 so that it could enhance the efficiency and will cost quite less; while being way faster in execution than the existing detection methods. This is exactly the main motive behind our deep learning model as it can be used as an automated diagnostic tool to predict COVID-19 from chest X-rays.

II. RELATED WORKS

Broad examination work is continuing for ordering COVID-19 patient chest x-ray information. Barely any scientists have proposed diverse DL models for arranging chest X-ray images while some others have contemplated CT images.

Narin et al. proposed 3 pre-trained CNN models dependent on ResNet50, InceptionV3, and Inception-ResNetV2 for distinguishing COVID-19 patients from chest X-ray radiographs. It is discovered that ResNet-50 gives a characterizing accuracy of 98% though InceptionV3 and Inception-ResNetV2 perform with the precision of 97% and 87% separately. In any case, these models have taken just 100 pictures (50 COVID19 and 50 typical Chest X-rays) into thought for preparing which may bring about declined exactness for a higher number of preparing images.

Zhang et al. propose a DL model for Coronavirus patient screening utilizing their chest X-ray images. This examination bunch has utilized 100 chest X-ray images of 70 COVID-19 patients and 1431 X-ray images of other pneumonia patients where they are delegated COVID-19 and non-COVID-19 separately. This model is shaped of three principle parts: spine organizations, grouping head, and peculiarity discovery head. The spine network is an 18 leftover CNN layer pre-prepared on the ImageNet dataset and it is mentionable that ImageNet gives an enormous summed up dataset to picture groupings. This model can analyze COVID-19 and non-COVID-19 patients with a precision of 96% and 70.65% separately.

Hall et al. likewise chipped away at discovering COVID-19 patients from a little arrangement of chest X-ray images with DL. They have utilized pre-prepared ResNet50 which creates the general exactness of 89.2%.

Sethy and Behea have likewise used profound highlights for Coronavirus illness identification. Their model depends on ResNet50 in addition to SVM which accomplished a precision of 95.38%.

Apostolpoulos and Mpesiana used CNN move learning for recognizing COVID-19 with X-ray images. This work has considered 224 chest X-ray images of COVID-19 contaminated individuals, 714 pictures with Pneumonia, and 504 pictures of typical individuals for preparing their model. This model accomplished a precision of 96.78%.

Islam et al. proposed a DL model dependent on CNN and long transient memory (LSTM). This work characterizes the dataset into three classifications—ordinary, COVID-19, and pneumonia with 1525 X-ray images in every one of those classifications. This CNN-LSTM based model accomplished a general precision of 99.4%.

S. No.	Model	Architecture	Non-COVID-19 Dataset	COVID-19 Dataset	Overall Accuracy (%)
1	Our Proposed Model	Sequential CNN	1420	1269	99
2	Deep Transfer Learning based Model by Loey et al. [58]	GoogleNet	69	79	99.9
3	Transfer Learning with CNN by Apostolopoulos and Mpesiana [33]	MobileNet V2	504	224	96.78
4	Transfer Learning Model by Sethy and Behera [32]	ResNet50+ SVM	25	25	95.38
5	COVIDX-Net by Hemdan et al. [36]	VGG19	25	25	90
6	Pre-trained CNN Model by Chowdhury et al. [44]	DenseNet201	1579	423	99.70
7	Deep Neural Networks by Ozturk et al. [45]	DarkCovidNet	500	127	98.08
8	A Deep CNN-LSTM Network by Islam et al. [35]	CNN-LSTM	1525	1525	99.4
9	A CNN Model by Haque et al. [53]	Sequential CNN	206	206	97.56
10	A Deep CNN Model by Narin et al. [29]	ResNet50	50	50	98
11	A Deep Learning Model by Zhang et al. [30]	Deep CNN	1431	100	96
12	Deep Learning Model by Hall et al. [31]	ResNet50	102	102	89.2
13	COVID-CAPS by Afshar et al. [49]	Capsule Networks	–	–	95.7
14	DeTraC Deep CNN by Abbas et al. [59]	DeTraC	80	105	95.12
15	COVID-CXNet by Haghanifar et al. [46]	Pre-trained CheXNet	3200	428	99.04
16	CoroNet by Khan et al. [47]	Xception	310	284	99
17	Deep Transfer Learning Model by Minaee et al. [48]	SqueezeNet	5000	184	92.29

Fig 1: Comparison of different Deep Learning Models based upon COVID X-Ray dataset and their performances

Hemdan et al. presented a profound learning structure naming AI 2020, 1 421 COVIDX-Net to arrange COVID-19 X-ray images. This model depends on just 25 chest X-ray images in every one of the classes—ordinary and Covid-19. For ordering the pictures this model uses seven distinctive pre-prepared models—VGG19, DenseNet121, InceptionV3, ResNetV2, Inception-ResNetV2, Xception, MobileNetV2. This exploration bunch has accomplished the best exhibitions from VGG19 with a general exactness of 90%.

Chowdhury et al. utilized exchange learning with picture expansion to distinguish COVID-19 from chest X-ray images. This work does the arrangement in two diverse plans—(I) COVID-19 and typical and (ii) COVID-19, viral pneumonia, and COVID-19. They have utilized 423 COVID-19, 1485 viral pneumonia, and 1579 typical chest X-ray images separately for preparation and approval. This gathering accomplished a fantastic outcome with paired grouping with a precision of 99.70%.

Ozturk et al. likewise does the parallel and multi-class arrangement with their proposed DarkCovidNet model which depends on the Darknet-19 model. This model has utilized 500 ordinary and 127 COVID-19 chest X-ray images for preparing and approval of their model. For parallel arrangement, this model has accomplished a normal general exactness of 98.08% though it is diminished to 87.02% if there should be an occurrence of multi-class order.

Abbas et al. utilized Decompose, Transfer, and Compose (DeTraC) which is a profound CNN for grouping COVID-19 positive pictures. It has utilized 80 ordinary and 105 COVID-19 chest X-ray images with a picture increase for preparing and approving the model. This model acquired the exactness of 95.12%.

Different scientists have likewise invested energy to recognize COVID-19 patients from CT checks. The majority of the talked-about research works in distinguishing COVID-19 use pre-trained models for their model design. These models are trained on more summed-up datasets like ImageNet. Here, a successive CNN model is recommended that is computationally productive because of its total straightforwardness in design, and this is trained from scratch with the relevant dataset. Besides, this model is prepared and approved with a relatively bigger dataset than the vast majority of different scientists to break down how the model performs with expansion in dataset and change in convolutional layers which is novel according to the examined writing.

III. PROPOSED MODEL

A. Create CNN model:

The dataset has been collected from two sources: GitHub and Kaggle, both of which are publicly available to download. It contains images of COVID-19 infected and Normal Chest X-rays which are distinguished into two sub-folders naming “covid” and “normal”. Both of these folders are originally stated in two classes “Train” and “Test”. The train data folder will be used to train our deep learning model, while the test data will be used to validate and evaluate results from the model.

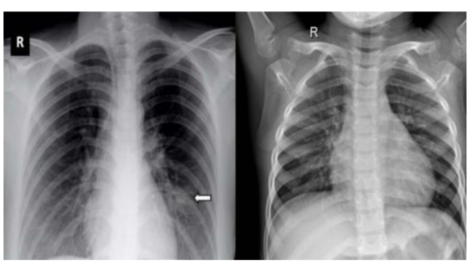

Fig 2: COVID-19 & Normal Chest X-Rays (LTR)

To create the CNN model, PyCharm IDE is used in which Python codes can be written and executed. To build the model, Numpy, matplotlib, Flask, CV2, Keras pre-defined libraries are used while TensorFlow is running in the backend. Before creating the model, image preprocessing is performed. The point of preprocessing is an improvement of the picture information that stifles undesired contortions or upgrades some picture highlights significant for additional handling and investigation task. The CNN model consists of two types of layers: Convolutional layers and Fully Connected Layers. A seven-layered CNN model is created, consisting of two sets of stacked convolutional and pooling layers followed by three fully connected layers for classifying COVID-19 and normal x-ray images. There are two convolutional layers with filter sizes of 3 x 3 over the layers. The standard kernel size is 3 x 3 and we have used the same. Activation of the ReLU function is used in the convolutional layers for non-linearity. Since this is the first layer, we specify the input size as 64 x 64 x 3. There are two pooling layers with filter size 2 x 2 each which is the default size, by using Max Pooling, the receptive field of the layer increases. The first convolutional layer is reshaped into 64 x 64 with 3 channels because the x-rays are RGB images. On carefully noticing the images, it is observed that the chest x-rays are not greyscaled images but RGB in nature because some pictures have a tone of blue or yellow. In the first set of a stacked convolutional layer, two convolutional layers of 3 x 3 kernel size have been used instead of one convolutional layer of 5 x 5 kernel size. Using two layers is advantageous to the model as it increases non-linearity in the model, and there are fewer parameters which in turn reduce overfitting. The model can recognize a more elevated level of features in images with the model layers extending. After every pooling layer, a dropout layer is added to reduce the risk of overfitting. After inputting the convolutional and pooling layers, the output shape has to be changed in order to go forward with the fully connected layers. Hence, the model is converted from a 2-dimensional layer to one dimension by using a flattening layer and then connected to a fully connected Dense layer. The fully connected Dense layer uses the ReLU activation function. There is only one filter applied for the output layer since we have to classify the images between COVID-19 and Normal chest x-ray images. Hence, the Sigmoid activation function is used. The model is compiled with binary entropy loss and Adam optimizer, which is the default optimizer function using accuracy metrics. The pre-defined Keras Image Generator library is used to train the dataset. In order to aid data convergence, the dataset is rescaled for normalization of RGB images by 1.0/255. Similarly, for the validation dataset, the Image Generator library is used again to rescale the images for normalization by 1.0/255. For our training dataset, the flow from the directory function is used to reshape the image with a target size of 64 x 64, having a batch size of 8, and using binary classification to distinguish between COVID-19 and Normal Chest x-rays. The equivalent has been done for our validation dataset. The model would be difficult to train if the input size would be big, and it would be challenging to capture fine-grained features if the input size would be too small. In the training process, 10 epochs are used with eight steps per epoch. This model is ridiculously simpler than other models like VGG16, ImageNet, and TransferNet as it uses few parameters. The other models will not give good results since they contain over a million parameters. Since the dataset contains chest x-rays, the models would have to be carefully fine-tuned with several modifications. Moreover, if the dataset is too small, and would hence lead to overfitting. The other models require at least

3000 images for training in order to give a good accuracy score but with a simpler approach, a lot of complexities have been ghosted.

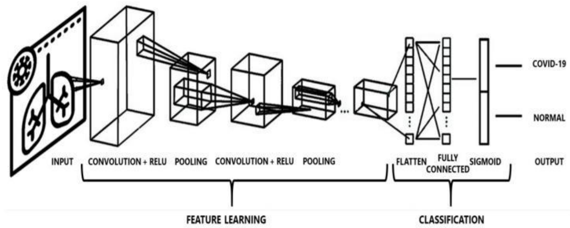

Fig 3: Overall workflow of our CNN model

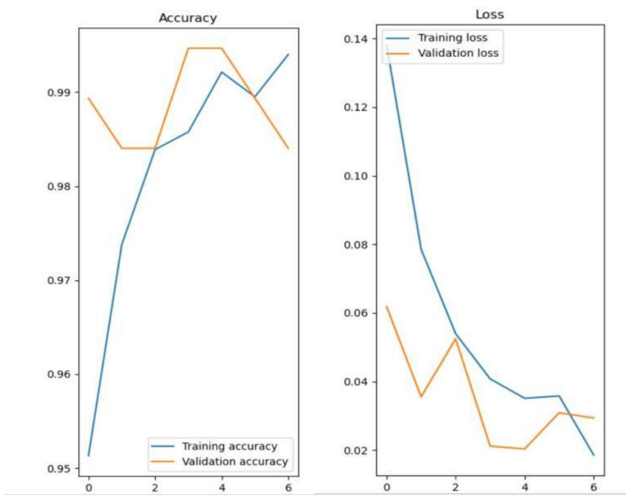

Fig 4: Accuracy and Loss graph for the trained model

Since the COVID-19 situation has created havoc for the whole world, our web app is combined with a database to store user records for the sake of medical emergencies. There are a lot of databases available such as MongoDB, MySQL, Cassandra, Google Firebase, etc. Google Firebase is Google-backed application software that enables developers to create apps on multiple platforms. It provides tools for database, hosting, authentication, storage, and many more. For our project, we have used its Realtime Database. The Firebase Realtime Database is a cloud-hosted NoSQL database that lets you store and sync data between your users in real-time. To get started with it, sign in to Google Firebase using a Google account. Then create a new Firebase project on its console. After this, we registered our web app to the Firebase project using firebase SDKs by connecting several APIs. And finally, initialize the firebase real-time database using initialize () method.

B. Create a web app:

The sole aim of this project is to discover a page where an actor can upload their chest X-Rays to get predictions of their COVID-19 status. For the creation of the front end of the web app, we have used HTML, CSS, JavaScript, Bootstrap technologies. Then we integrated the earlier saved model in the UI/UX using Flask. Flask is a micro web framework written in Python. It has no database abstraction layer, form validation, or other components. However, it supports extensions that can add application features as if they were implemented in Flask itself, just like we did with our deep learning model. We rendered the website files in Flask and created a user-defined method to evaluate results from our existing model which is called using load model (). In this function, the image from the user will be preprocessed using OpenCV and Numpy and then will be saved in the local directory. This saved image will then be passed through our intelligent deep learning model using the model. predict () method. The output of this function is passed as a reference in a variable called “classes” which is again passed on to another variable called “label” such that the result keeps itself clean from garbage values. This “label" is then checked under a condition for its binary mode existence, i.e., if the label gives the output as 0, the image provided is assumed as COVID-19 positive, otherwise, it is assumed as Normal.

C. Setting up the application's database:

Since the COVID-19 situation has created havoc for the whole world, our web app is combined with a database to store user records for the sake of medical emergencies. There are a lot of databases available such as MongoDB, MySQL, Cassandra, Google Firebase, etc. Google Firebase is Google-backed application software that enables developers to create apps on multiple platforms. It provides tools for database, hosting, authentication, storage, and many more. For our project, we have used its Realtime Database. The Firebase Realtime Database is a cloud-hosted NoSQL database that lets you store and sync data between your users in real-time. To get started with it, sign in to Google Firebase using a Google account. Then create a new Firebase project on its console. After this, we registered our web app to the Firebase project using firebase SDKs by connecting several APIs. And finally, initialize the firebase real-time database using initialize () method.

IV. RESULTS & ANALYSIS

In order to get a summary of the results on a classification problem, a confusion matrix can be used to compute and visually understand the expected results of the model. It can help to summarize the correct and incorrect predictions broken down by each class. It shows how the classification model gets confused when it makes predictions and gives an insight not only into the magnitude of errors being made but also regarding the type of errors being made.

	COVID-19	Normal
COVID-19	True Positive	False Negative
Normal	False Positive	True Negative

Where True Positive: Observation is positive, the prediction is positive, False Negative: Observation is positive, the prediction is negative, False Positive: Observation is negative, the prediction is positive and True Negative: Observation is negative, the prediction is negative.

V. CONCLUSION

Mass testing and the early location of COVID-19 assume a significant part in forestalling the spread of this new worldwide pandemic. Time, cost, and precision are a couple of main considerations in any illness location measure uniquely COVID-19. To address these issues, a CNN-based model is proposed in this project report for detecting COVID19 cases from patients' chest X-rays. The CNN model is trained with Dataset which has a total of 2689 chest X-ray images divided into two classes. The proposed model excels with an accuracy of 99 percent. Moreover, this model compares the achieved results with other prominent works in the field by different people and found it to better than most of them all. This work can be improved further with the availability of a larger dataset. Finally, CNN has extraordinary possibilities in identifying COVID-19 with exceptionally restricted time, assets, and expenses. Despite the fact that the proposed model shows promising outcomes, it's anything but clinically tried. However, with such a higher accuracy the planned model will sure play a crucial role in the early and quick detection of COVID-19, therefore, reducing testing time and price. Early prediction of COVID-19 patients is significant to forestall the unfold of the sickness to people. The virus is generally new, and no authority immunization has been begun at this point. Hence, humanity ought to find different ways to prevent the spread of COVID-19 as soon as possible. Nonetheless, the present work contributes to the possibility of a low-cost, rapid, and automatic diagnosis of COVID-19. Also, even though the appropriate treatment is not determined solely from an X-ray image, an initial screening of the cases would be useful, not in the type of treatment, yet, in the convenient use of isolate measures in the positive examples, until a total assessment and explicit treatment or follow-up system are followed. An additional advantage of automatic detection of COVID-19 from medical imaging lies in the reduction of exposure of nursing and medical staff to the outbreak.

REFERENCES

[1] Haque, K. and Abdelgawad, A., 2020. A Deep Learning Approach to Detect COVID-19 Patients from Chest X-ray Images. *AI*, 1(3), pp.418-435.

[2] Haritha, D., Krishna Pranathi, M., and Reethika, M., 2021. COVID Detection from Chest X-rays with DeepLearning: CheXNet. *IEEE*.

[3] Jain, R., Gupta, M., Taneja, S., and Hemanth, D., 2020. Deep learning-based detection and analysis of COVID-19 on chest X-ray images. *Applied Intelligence*, 51(3), pp.1690-1700.

[4] Boudrioua, M., 2020. COVID-19 Detection from Chest X-Ray Images Using CNNs Models: Further Evidence from Deep Transfer Learning. *SSRN Electronic Journal*.

[5] T. Padma and C. U. Kumari, "Deep Learning-Based Chest X-Ray Image as a Diagnostic Tool for COVID-19," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 589-592, DOI: 10.1109/ICOSEC49089.2020.9215257.

[6] SHARMA, S., 2021. *Activation Functions in Neural Networks*. [online] Medium. Available at: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6> [Accessed 28 June 2021].

[7] Godoy, D., 2021. *Understanding binary cross-entropy / log loss: a visual explanation*. [online] Medium. Available at: <https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a> [Accessed 28 June 2021].

[8] En.wikipedia.org. 2021. *Flask (web framework) - Wikipedia*. [online] Available at: <https://en.wikipedia.org/wiki/Flask_(web_framework)> [Accessed 28 June 2021].

[9] GeeksforGeeks. 2021. *Keras.Conv2D Class - GeeksforGeeks*. [online] Available at: <https://www.geeksforgeeks.org/keras-conv2d-class/> [Accessed 28 June 2021].

[10] Brownlee, J., 2021. *A Gentle Introduction to Pooling Layers for Convolutional Neural Networks*. [online] Machine Learning Mastery. Available at: <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>