

Automatic Detection and Classification of Corona Infection (COVID-19) from X-Ray Images Using Convolution Neural Network

^[1] Kinjal A. Patel, ^[2] Tanvi Goswami

^[1] Department of Computer Application and IT, Gujarat Law Society University, Ahmedabad, India ^[2] Department of Information Technology, Dharmsinh Desai University, Nadiad, India Email: ^[1] kinjal5721@gmail.com, ^[2] Goswami.tanvi@gmail.com

Abstract---The novel coronavirus universally known as the COVID-19 outbreak arises at the end of 2019 in one of the East Asian countries and it is subjected to widespread discussion and debate. There are almost 200 countries affected across the globe by COVID-19. And, it has ruined many lives and the global economy. The virus is spreading very rapidly at the pace of around 10 fold in less than a month. Also, in the case of COVID- 19, it is critical to detect the infection, as it employs various symptoms which may differ from person to person. Hence, diagnosis in starting stage and treatment are very much important for such type of infectious disease. The chest x-ray is one of the primary techniques among blood tests and Computed Tomography contributes a major role in the early diagnosis of COVID-19. There is a rising need for automated and auxiliary diagnostic tools for early diagnosis, as there are no accurate and truthful automated toolkits on hand. In this research study, we have designed a Convolution Neural Network architecture – a deep net for the classification of x-ray images of chest among two classes: COVID-19 or Non-COVID-19 infection. The anticipated model is expected to provide accurate diagnostic results and produced classification accuracy of 99%, 100%, and 100% with 70%-30%,75%-25% and 80%-20% train-test data split respectively, for the binary classification of the x-ray image to be COVID-19 or Non-COVID-19 infection category. We have designed the CNN with optimized parameters with 3 convolution layers and optimized number of filters in each layer.

Key Words--- Deep Learning, COVID-19, Chest X-ray, Convolutional Neural Network, Instinctive detection

I. INTRODUCTION

The coronaviruses (CoVs) are most infectious in birds but, from several decades it is capable to change host and infecting the humans as well. The coronavirus also known as SARS-CoV-2 is a new member of SARS (severe acute respiratory syndrome) virus spices recognized by its genome sequences [1]. The novel coronavirus epidemic arises in one of the East Asian countries during December 2019 and has extensively blowout across the countries. The coronavirus is rapidly spread from humans to humans. The novel coronavirus indications vary from one person to another person. The extreme symptoms of the COVID-19 are high body temperature, dry cough, and weariness as well as few symptoms are sore throat, diarrhea, and headache [2]. The serious symptoms of the COVID-19 are chest aching and struggle in breath [2]. There are 7818 coronavirus positive cases and 170 deaths are confirmed across the globe during the end of January [3]. The "Public Health Emergency of International Concern" has been declared by The World Health Organization (WHO) on Janyary 30, 2020 due to the extensive increment in coronavirus positive cases [4]. WHO entitled this disease as COVID-19 by February 2020 [5]. As

of 25 May 2020, 53,04,772 COVID-19 positive cases and 3,42,029 deaths are confirmed across the globe and 1,50,762 COVID-19 positive cases, and 4,349 deaths are confirmed across India [6] [7].

The COVID-19 is transmissible, thus many countries are applied lockdown to prevent the spreading of COVID-19. Due to this kind of situation, the Government of India has launched many competitions to get innovative ideas from intelligent minds to deal with every domain such as patient care management, fake news detection, movement tracking, large scale sterilization, stabilizing affected business, virus containment and video conferencing for online education [8]-[11].

The numbers of researchers are working on finding proper treatment, antivirus drugs, and therapeutic vaccines of COVID-19. The detection of COVID-19 infection at a premature stage and isolation of the disease infected people is mandatory as the antivirus drug and therapeutic vaccine was not available. The unexpected hike in a positive COVID-19 leads to the need for a massive amount of medical resources. The real-time reverse transcriptionpolymerase chain reaction (RT-PCR) test is used to detect COVID-19 infection. But, RT-PCR test takes more time to



give the result as well as the result is not precise. RT-PCR has high false-negative rates and low sensitivity [12] [16]. The less accurate RT-PCR is not adequate in the present epidemic condition. In many cases, the negative results are acquired from the RT-PCR test, and COVID-19 infected person may not be identified, thus the healthy people may get infected from COVID-19 positive people. The indication of COVID-19 may include dry cough and chest pain and it can be examined using radiological imaging of the lungs and chest, computed tomography (CT) and X-ray. The CT has high sensitivity and accurate results [12]. As the CT is more accurate therefore it is useful to detect the infection of COVID-19 at an early stage. The radiologists examine the CT and X-ray images. Due to the hike in COVID-19 positive cases and shortage of medical professionals as well as RT-PCR tool kits, there is a requirement for the development of an automatic tool that detects the COVID-19 infectivity from CT scans and X- ray imagery. The deep learning based method come up to classify the COVID-19 and Non-COVID-19 infections from CT and X-ray images is worth advantageous while a shortage of medical professionals.

The core objective of this paper is to classify COVID-19 and Non-COVID-19 diseased person from X-ray images of the chest. The deep learning approach is used to categorize infected COVID-19 patients and non-infected COVID-19. The 2 dimensional convolution neural network (CNN) is designed for binary classification which classifies the X-ray images into two classes namely; COVID-19 and Non COVID-19. The CNN model is implemented with three 2D - convolutional layers and trained by chest X-ray images. The model performance is evaluated with different evaluation measures such as precision, recall, f-score, accuracy, and confusion matrices.

The outline of the paper is as follows: the work has been completed in the domain of COVID-19 is discusses in "Literature review" section; the implemented CNN model for classification is discussed in the "Proposed CNN model" section; the performance of CNN model and its comparison with existing models is discussed in the "Performance comparison and analysis" section; the paper is accomplished in the "Conclusion" section.

II. LIETRATURE REVIEW

The researchers have discovered the identity of the COVID-19 from patterns on images of chest CT and X-ray [16]-[24]. Due to the shortage and less sensitivity of RT-PCR kits, various comparative surveys are conducted with RT-PCR test and chest CT and X-ray. The study was conducted by Xie et al. [14] to identify COVID-19 infectious from 167 people. A total of 3% of 167 people had negative RT- PCR for COVID-19 infectious. But, that 3% of people were founded positive COVID-19 infectious on chest CT. The other study was conducted by Berheim et. al [15] by examining 121 chest CTs of COVID-19 infected patients. The increased severity of infection is founded in concerning time from initial infection was detected. Fang et al. [13] examined the 2 patients along with their symptoms as well as travel history. They came up with the results that CT of chest had a higher sensitivity as compared to RT-PCR to identify COVID-19.

Ozturk et al. [16] developed the DarkCovidNet model which has 17 convolutional layers with different filters at every layer. They used DarkCovidNet for 2 classes namely; for COVID and No-findings as well as the multi-class classification for Pneumonia, COVID, and No- findings. The model formed 98.08% of classification accuracy for 2class classification and 87.02% of classification accuracy for multi-class classification. Singh et. al [17] developed a CNN to classify COVID-19 positive infection or not from CT images of chest. They tuned the parameters of CNN by using multi-objective differential evolution (MODE). The MODE based CNN gives a significant performance with 1.9789% accuracy, 2.0928% F-measure, 1.8262%

sensitivity, 1.6827% specificity, and 1.9276% Kappa statistics.

Khan et al. [18] extract the features from x- ray images of the chest of COVID-19 patients and fed that features into the SVM for classification of positive COVID-19 or negative COVID-19. The developed COVID-RENet and Custom VGG model give accuracy 98.3%, AUC 0.98, Fscore 0.98, Recall 96.67%, and Precision 100%. Wang et. al. [19] proposed CNN as a COVID-Net which as lower complexity of architecture and computation as compare to VGG-19 and ResNet-50. The COVID-Net was already trained on the ImageNet and then again trained that model on the COVIDX dataset to categorize Pneumonia, COVID-19, and normal patients. The 93.3% classification accuracy was obtained from custom developed model COVID-net and on the other hand, 83.0% classification accuracy obtained from VGG-19 and 90.6% classification accuracy obtained from ResNet-50.

Narin et al. [20] have used ResNet50, InceptionV3, and Inception-ResNetV2 to categorize between normal X-ray images and COVID-19 infected images. To overcome the shortage of sufficient data, they used the ImageNet dataset as well. Also, transfer learning technique was employed which also reduces the time needed for training the model. The 98% of classification accuracy is obtained by pertained ResNet-50 model whereas 97% and 87% of classification accuracy obtained by InceptionV3 model and Inception-ResNetV2 respectively. Hasan et al [21] has projected a Q-



deformed entropy algorithm based on deep learning for feature extraction from lung images. These extracted features fed into LSTM) neural network which classifies: Pneumonia, COVID- 19 and healthy lungs. They obtained 99.68% accuracy from the projected model.

Apostolopoulos et al. [22] has used various existing pretrained CNN architecture which are: MobileNet V2, VGG 19, Xception, Inception Resnet v2 and Inception. They come up with the comparative study of these models and applied transfer learning to classify viral and bacterial pneumonia, COVID-19, and also normal lungs conditions from X-ray images. There study includes two datasets. Dataset-1 with two classes' normal and COVID-19 and Dataset-2 with three classes normal conditions, viral pneumonia and COVID-19. The highest accuracy is obtained in VGG-19 which is 98.75% for dataset-1 and 93.48% for dataset-2 among all the models. The MobileNet V2 obtained 97.40% and 92.85% accuracy, Inception obtained 86.13% and 92.85% accuracy, Xception obtained 85.57% and 92.85% accuracy and Inception Resnet v2 obtained 84.38% and 92.85% accuracy for dataset-1 and dataset-2 in that order. Abbas et al. [23] uses customdeveloped CNN - DeTrac known for Decompose, Transfer, and Compose to classify the COVID from normal cases and SARS cases by using X-ray images. The concept of DeTrac handles the irregular image dataset. They obtained accuracy of 95.12%, sensitivity 97.91%, specificity 91.87%, and precision 93.36%. Khalid et al. [24] has come up with comparative analysis of various CNN pre-trained architectures which are: DenseNet201, VGG16, Inception ResNet V2, VGG19, Resnet50, MobileNet V2 and InceptionV3 that detect and classify normal, COVID-19 and bacterial pneumonia chest X-ray images of chest and CT scan images. They obtained best accuracy of 92.18% from and 88.09% with Densnet201.

From this comprehensive analysis of existing studies, it is concluded that the CT images and X-ray images is likely to be considered for detecting and classifying the infection at its early stage [25]. Hence, in this paper, Convolutional Neural Network – deep learning model is implemented to classify COVID-19 patients from chest X-ray images.

III. PROPOSED WORK

A. Dataset Collection

Data is the primary step for developing any computer-aided diagnostic tool. The dataset for this experiment was gathered from [26]. The Dataset is having 912 images for each class. These images are original x-ray images as well as augmented images for two classes namely: COVID-19 and Non-COVID-19 detection from X-ray images of chest. COVID-19 images are original and augmented X-ray

images whereas non-COVID-19 images are original images [26]. The sample image from each class is demonstrated in Fig. 1.



Fig. 1 (a) COVID-19 image, (b) Non-COVID-19 image [26]

B. Proposed CNN Architecture

We have observed and studied various parameters and hyper-parameters of CNN that is; Number of convolutional filters, number of convolutional layers, kernel size, pooling window size, fully connected layers, and number of neurons in fully connected layers. These parameters are selected after performing an analysis of works carried out by observing available pre-trained architectures used by researchers so far [19] [20]. As an outcome of the analysis, we have designed CNN architecture from scratch by keeping all the parameters of CNN into consideration.



The proposed architecture of CNN is presented in Fig. 2. Our proposed model consists of three convolutional layers followed by 3 max-pooling layer, 2 fully connected layers, and an output layer consist of softmax activation fuction. We evaluated our model on the augmented COVID-19 Xray image dataset [26]. We split the dataset into three train and test sets: 70% - 30%, 75% - 25% and 80% - 20% traintest split. The model was trained for 45 epochs. The default input image size of 256×256 was considered in the proposed model evaluation. Resizing is performed to change the various image dimensions to 256×256 . The optimizer used was Adam 0.001learning rate throughout the experiments because Adam performs better than Gradient Descent in many machine learning applications. Drop out



method was also used to avoid overfitting, which randomly drops connections from the previous layer. The values of Adam's learning rate and the number of epochs were selected after certain observations of training loss. The architecture details are listed in Table 1.

-										
Layer	Processing unit	Input Dimension	Output Dimension	Kernel Size	No. of Kernels	Zero padding				
1	Conv2D	256x256	256x256x64	3x3	64	2x2				
2	MaxPool	256x256x64	130x130x64	2x2	64	-				
3	Conv2D	130x130x64	130x130x128	3x3	128	2x2				
4	MaxPool	130x130x128	67x67x128	2x2	128	-				
5	Conv2D	67x67x128	67x67x128	3x3	128	2x2				
6	MaxPool	67x67x128	35x35x128	2x2	128	-				
7	Fully Connected Layer 1: 1024 neurons followed by activation ReLu and 0.2 DropOut rate									
8	Fully Connected Layer 2 :1024 neurons followed by activation ReLu and 0.2 DropOut rate									
9	Output Layer with softmax activation function									

TABLE I: CONFIGURATIONS OF PROPOSED THE CNN ARCHITECTURE

To avoid shrinkage of dimensions, zero-padding with size 2x2 is used. Steps per epoch are also one of the hyperparameters of CNN architecture; its value can be determined by batch size as well as the number of training samples. A large batch size provides a better gradient and prevents jumping around, but too large batch size may create memory problems. After analyzing the available memory we kept 64 batch sizes. Also, we kept the kernel size 3x3 across the experiments.

IV. RESULTS AND DISCUSSION

A. Performance Evaluation

The dataset was divided into three different splits; 70% - 30%, 75% - 25% and 80% - 20% train- test splits respectively. We have designed the CNN architecture as shown in Fig. 2. The model is trained for 45 epoch. The results are summarized in Table 2.

The confusion matrix are shown in Fig. 3 with COVID-19 and Non-COVID-19 test results of the designed architecture.



Fig. 3 The confusion matrix was obtained for three different splits using the proposed CNN model. The results for three experiments are shown in: a) confusion matrix for 70%-30% split b) confusion matrix for 75%-25% split, c) confusion matrix for 80%-20% split

We have obtained the best performance with 100% accuracy, 100% recall, and 100% precision for 75%-25% and 80%-20% splits. The lowest performance values have been obtained in 70%-30% splits with 99% accuracy, 99%

recall and 99% precision. As a result experiment number 2 and 3 as mentioned in Table 2 provides better results as compared to experiment number 1.



TABLE II: PREDICTION PERFORMANCE RESULTS FROM THE CNN MODEL WITH DIFFERENT SPLITS												
Experiment	Train-test Split	No. of training	No. of Testing	Confusion Matrix and Performance Result		ce Results in						
Number		images	Images	% over the test set								
				Precision	Recall	F1 Score	Accuracy					
1	70%-30%	1282	542	99	99	99	99					
2	75%-25%	1370	454	100	100	100	100					
3	80%-20%	1460	364	100	100	100	100					

V. DISCUSSION

For this research work, the experiments were performed using Google Colaboratory with Tesla K80 GPU for Tensorflow and Keras library. The COVID-19 dataset is very limited also few research works have been reported as it is new to everyone so far. The major advantages of the proposed CNN model are: - 1) The model aims to classify chest X-ray images without using the traditional methods. The traditional technique uses different feature extraction and segmentation methods additionally while CNN extracts the features automatically. 2) The automated the approach is very much useful in assisting experts in decision making.

VI. CONCLUSION

In this paper, a COVID-19 and Non-COVID-19 infection classification model is implemented to classify the infected humans from chest X-ray images. Primarily, the chest X-ray dataset of COVID-19 and Non-COVID-19 is decomposed into 70%- 30%, 75%-25%, and 80%-20% train-test data split. The data from each train-test split is applied separately to the training of the CNN model. The CNN model is trained with 45 epochs for each train- test data split. As a final point, the comparisons are drawn from each CNN model by considering diverse fractions of training- testing dataset. All-inclusive experimental results disclose that the proposed model outperforms in terms of accuracy, precision, recall, F1-score by 99% for each in 70%-30% split and 100% for each in the remaining two data splits. Hence, the proposed model is beneficial for instantaneous COVID-19 disease classification from chest X-ray images.

REFERENCES:

- Y., Ho, W., Huang, Y., Jin, DY, Li, S., Liu, SL, Liu, X., Qiu, J., Sang, Y., Wang, Q., Yuen, KY, & Zheng, ZM Wu, "SARS-CoV-2 apt name for new coronavirus," Lancet (London, England), 395 (10228), vol. https://doi.org/10.1016/S0140-6736 (20) 30557-2, pages 949-950, 2020.
- [2] (2020, May) World Health Organization WHO Health. [Online]. HYPERLINK "https://www.who.int/health-topics/coronavirus" \ 1 "tab = tab_3" https://www.who.int/healthtopics/coronavirus#tab=tab_3

- [3] (2020, January) World Health Organization WHO Status-Daily Reports. [Online]. HYPERLINK "https://www.who.int/docs/defaultsource/coronaviruse/situation-reports/20200130-sitrep-10-ncov.pdf?sfvrsn=d0b2e480_2" https://www.who.int/ docs / default-source / coronaviruse / status reports / 20200130-sitrep-10ncov.pdf? sfvrsn = d0b2e480_2
 [4] (2020, January) World Health Organization, WHO
- [4] (2020, January) World Health Organization WHO newsroom details. [Online]. HYPERLINK "https://www.who.int/news-room/detail/30-01-2020statement-on-the-second-meeting-of-the-internationalhealth-regulations-(2005)-emergencythe-novel-coronavirus- (2019-ncov) "https://www.who.int/news-room/detail/30-01-2020statement-on-the-second- health-law-meeting-health -International Committee- (2005-ncov)
- [5] (2019) World Health Organization COVID-19 Technical Guide. [Online]. HYPERLINK "https://www.who.int/emergency/diseases/novelcoronavirus-2019/technical-guidance/naming-thecoronavirus-disease-(covid-2019)-and-the-virus-thatcauses- it
 "https://www.who.int/emergency/diseases/novelcoronavirus-2019/technical-guidance/naming-thecoronavirus-2019/technical-guidance/naming-thecoronavirus-disease-(covid-2019)-and-the-virus-thatcauses- it
- [6] (2020, May) World Health Organization WHO Status
 Daily Reports. [Online]. HYPERLINK "https://www.who.int/docs/defaultsource/coronaviruse/situation-reports/20200525-covid-19-sitrep-126.pdf?sfvrsn=887dbd66_2" https: //www.who. int / documents / default-source / coronaviruse / status reports / 20200525-covid-19sitrep-126.pdf? sfvrsn = 887dbd66_2
- [7] (2020, May) WorldOmeter. [Online]. https://www.worldometers.info/coronavirus/country/in dia/
- [8] (2020, May) The Challenge of Designing a Video Conference Solution. [Online]. https://startups.meitystartuphub.in/public/application/i nc/5e92ec1269e3401cd7bc6db7
- [9] (2020, May) Launch of India. [Online]. https://www.startupindia.gov.in/content/sih/en/ams-



application/challenge.html?applicationId=5e79126ee4 b055bfaea9ef66

- [10] (2020, May) MyGOV COVID-19 Solution Challenge. [Online]. https://www.mygov.in/task/covid-19solution-challenge/
- [11] (2020, May) Student Open Innovation Challenge. [Online]. http://www.ssipgujarat.in/soic/index.php
- [12] Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen, Wenzhi Lv3, Qian Tao, Ziyong Sun, Liming Xia Tao Ai, "Correlation of Chest CT and RT-PCR Testing in Coronavirus," Radiology, no. 200642. Prior to publication online. https://doi.org/10.1148/radiol.2020200642, February 2020.
- [13] Y., Zhang, H., Xu, Y., Xie, J., Pang, P., & Ji, W Fang, "CT Manifestations of Two Cases of 2019 Novel Coronavirus (2019-nCoV) Pneumonia," Radiology, vol. 295 (1), no. https://doi.org/10.1148/radiol.2020200280, pages 208-209, April 2019.
- [14] Zhong Z, Zhao W, Zheng C, Wang F, Liu J Xie X, "Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing," Radiology, no. doi: 10.1148 / radiol.2020200343, Feb 2020; 200343.
- [15] Adam and May, Xueyan and Huang, Mingqian and Yang, Yang and Fayad, Zahi and Zhang, Ning and Diao, Kaiyue and Lin, Bin and Zhu, Xiqi and Li, Kunwei and Li, Shaolin and Shan, Hong and Jacobi, Adam & Chung, Michael Bernheim, "Chest CT Detection in Coronavirus Disease-19 (COVID-19): Relationships and Time of Infection, "Radiology, vol. 295. 200463., no. DOI: 10.1148 / radiol.2020200463, February 2020.
- [16] Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, U. Rajendra Acharya Tulin Ozturk, "Automatic detection of COVID-19 cases using four deep neural networks," Computers in Biology and Medicine 121 (2020) 103792, vol. 121, no.

https://doi.org/10.1016/j.compbiomed.2020.103792, June 2020.

- [17] Vijay Kumar, Vaishali, Manjit Kaur Dilbag Singh, "Separating COVID-19 patients from chest CT images using a variety of variables based on convolutional neural networks," European Journal of Clinical Microbiology & Infectious Diseases, no. https://doi.org/10.1007/s10096-020-03901-z, April 2020.
- [18] Saddam & Sohail, Anabia & Zafar, Muhammad & Khan, Asifullah Hussain, "Coronavirus Disease Analysis using Chest X-ray Images and Novel Deep Convolutional Neural Network," 2020.

- [19] Alexander Wong Linda Wang, "COVID-Net: A Designed Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," Electrical Engineering and Systems Science> Image and Video Processing, no. arXiv.org> eess> arXiv: 2003.09871, May 2020.
- [20] Ceren Kaya, Ziynet Pamuk Ali Narin, "Automatic Detection of Coronavirus (COVID-19) Using X-ray Imaging and Deep Convolutional Neural Networks," Electrical Engineering and Systems Science, Image and Video Processing, no. arXiv.org> eess> arXiv: 2003.10849, March 2020.
- [21] Ali & Al-Jawad, Mohammed & Jalab, Hamid & Shaiba, Hadil & Ibrahim, Rabha & Shamasneh, Alaa Hasan, "classification of Covid-19 Coronavirus, Pneumonia and Healthy Lungs in CT Scans Using Q-Deformed Entropy and Deep Learning Features, "Entropy DOI: 10.3390 / e22050517, vol. 22, no. 5, page. 2020, April 2020.
- [22] I.D., Mpesiana, T.A. Apostolopoulos, "Covid-19: Automatic Discovery from X-ray images used to transfer learning through convolutional neural networks," Phys Eng Sci Med (2020), no. https://doi.org/10.1007/s13246-020-00865-4, March 2020.
- [23] Mohammed Abdelsamea, Mohamed Gaber Asmaa Abbas, "Separation of COVID-19 in chest X-ray images using DeTraC deep network convolutional neural network," medRxiv - The Preprint Server For Health Sciences, no. https://doi.org/10.1101/2020.03.30.20047456, May 2020.
- [24] Chawki and Elasnaoui, Khalid Youness, "Use X-ray Images and in-depth study to detect automatic diagnosis of Coronavirus," Journal of Biomolecular Structure & Dynamics, no. SECTION: 10.1080 / 07391102.2020.1767212., May 2020.
- [25] Xu P, Lv WF, Qiu XH, Yao JL, Jin-Feng G Liu KC, "CT manifestations of coronavirus disease-2019: a retrospective analysis of 73 cases by severe disease," European Journal of Radiology 126: 108941, no. DOI: 10.1016 / j.ejrad.2020.108941, March 2020.
- [26] Shoroq Qazan Ali Mohammad Alqudah. (2020, March) Progress. [Online]. HYPERLINK https://data.mendeley.com/datasets/2fxz4px6d8/4" \ 1 "folder-2eec69c7-9b6c-4864-b9d5-bd059e2dd24b.