

Data science and the role of Artificial Intelligence in achieving the fast diagnosis of Covid-19

Dasari Naga Vinod^a, S.R.S. Prabaharan^{b,*}

^a Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Potheri, Chengalpattu District, Tamilnadu, India

^b Directorate of Research, SRM Institute of Science and Technology, Potheri, Chengalpattu District, Tamilnadu, 603203, India

ARTICLE INFO

Article history:

Received 9 May 2020

Accepted 29 July 2020

Available online 30 July 2020

Keywords:

Data Science

Covid-19

Computed tomography

Decision tree classifier

X-ray images

Artificial Intelligence

ABSTRACT

The rapid spread of novel coronavirus (namely Covid-19) worldwide has alarmed a pandemic since its outbreak in the city of Wuhan, China in December 2019. While the world still tries to wrap its head around as to how to contain the rapid spread of the novel coronavirus, the pandemic has already claimed several thousand lives throughout the world. Yet, the diagnosis of virus spread in humans has proven complexity. A blend of computed tomography imaging, entire genome sequencing, and electron microscopy have been at first adapted to screen and distinguish SARS-CoV-2, the viral etiology of Covid-19. There are a less number of Covid-19 test kits accessible in hospitals because of the expanding cases every day. Accordingly, it is required to utensil a self-exposure framework as a fast substitute analysis to contain Covid-19 spreading among individuals considering the world at large. In the present work, we have elaborated a prudent methodology that helps identify Covid-19 infected people among the normal individuals by utilizing CT scan and chest x-ray images using Artificial Intelligence (AI). The strategy works with a dataset of Covid-19 and normal chest x-ray images. The image diagnosis tool utilizes decision tree classifier for finding novel corona virus infected person. The percentage accuracy of an image is analyzed in terms of precision, recall score and F1 score. The outcome depends on the information accessible in the store of Kaggle and Open-I according to their approved chest X-ray and CT scan images. Interestingly, the test methodology demonstrates that the intended algorithm is robust, accurate and precise. Our technique accomplishes the exactness focused on the AI innovation which provides faster results during both training and inference.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

2019 version of "Coronavirus" is often represented with "novel", as it is another strain in the group of infections that we have all watched already. According to WHO's classification, Covid-19 virus strains have a place with a huge family which runs from the normal cold to hazardous ailments [1]. These sicknesses can blemish both in human and animals. The strain that began spreading in Wuhan in late 2019, the capital of China's Hubei province, is distinguished from two distinctive coronaviruses, SARS and MERS. The vast majority falls debilitated with Covid-19 will encounter gentle to direct side effects and recoup without unique treatment [2].

Let us see how it spreads. The infection that causes Covid-19 is principally transmitted through droplets created when an infected individual coughs, sneezes, or breathes out. These droplets are too

substantial to even think about hanging noticeable all around, and rapidly fall on floors or surfaces. People can be infected by breathing in the infection in the event that individuals are inside closeness of somebody who has Covid-19, or by contacting a polluted surface and afterward your eyes, nose or mouth. Covid-19 influences various individuals in various manners. Most contaminated individuals will create gentle to direct side effects. Some prominent symptoms include fever, fatigue, breathlessness, dry cough, and hiccups. Some people may experience aches and pains, nasal congestion, runny nose, sore throat, diarrhoea [3].

During normal course, it takes 5–6 days from when somebody is contaminated with the infection for side effects to appear, however, it can take as long as 14 to 21 days. Individuals with mild side effects who are in any case healthy should self-isolate themselves. Perhaps, depending upon individuals conditions such as fever, cough, and difficulty breathing.

The analysis of Covid-19 is as of now a troublesome errand on account of inaccessibility of analysis framework all over the

* Corresponding author.

E-mail address: srsprabaharan1611@gmail.com (S.R.S. Prabaharan).

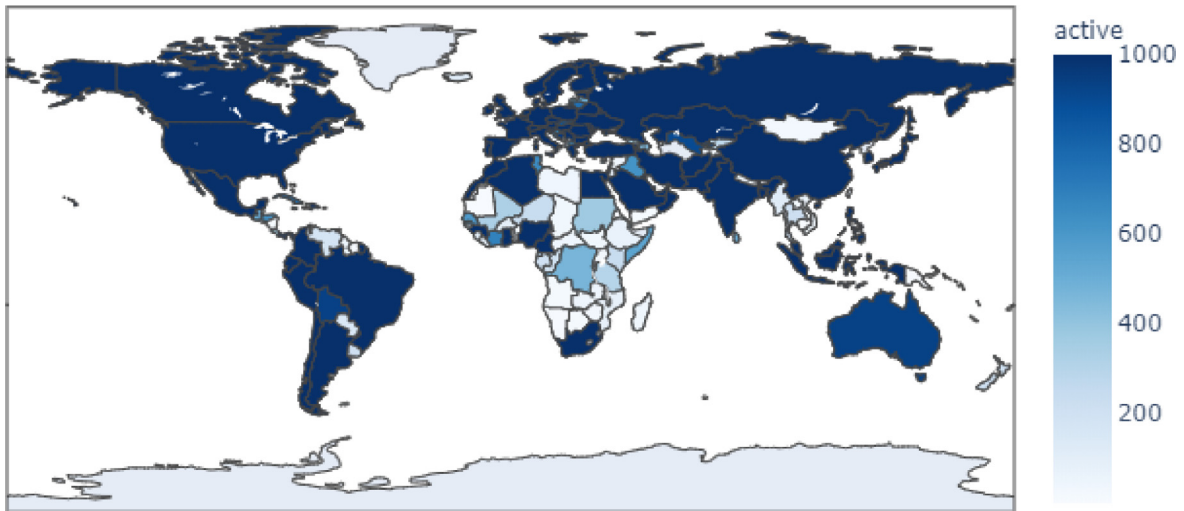


Fig. 1. Shows active cases worldwide.

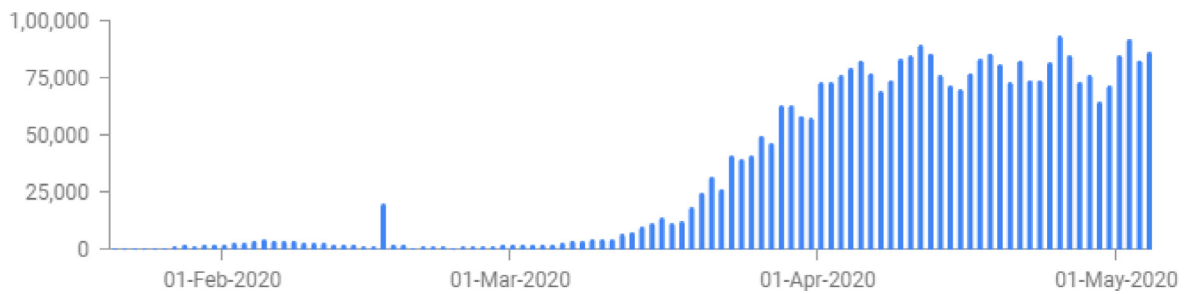


Fig. 2. Graph shows new cases over time (worldwide) [5].

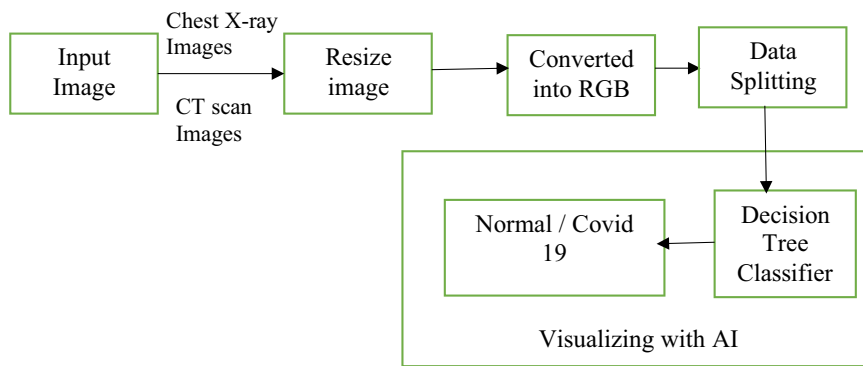


Fig. 3. Block diagram shows recommended approach.

place, which is causing fear. Due to the restricted accessibility of Covid-19 testing kits, we need to depend on different conclusion measures. Since Covid-19 assaults the epithelial cells that line our respiratory tract, we can utilize X-ray and CT images to investigate the soundness of a patient’s lungs. The clinical specialists every now and then utilizes X-ray and CT scan pictures to analyze pneumonia, lung aggravation, abscesses, as well as developed lymph hubs [4]. Furthermore, nearly in all medical clinics have X-beam imaging machines, it could be conceivable to utilize X- ray and CT to test for Covid-19 without the committed test packs. Moreover, such routine imaging techniques poses disadvantages that X-ray investigation requires a radiology master and takes critical time, which is valuable when individuals are wiped out around the world. Along these lines, building up a robotized investigation framework is important to spare clinical experts precious time.

As of now, the world is under utmost danger owing to Covid-19 spread by an infection of novel corona virus family. Most part of the world nations has seen countless Covid-19 cases since December 2019 onwards. Active cases of individuals with low resistance, mature age, and clinical issues extraordinarily identified with Covid-19 related lung infections that are progressively inclined as shown in Fig. 1.

Fig. 2 shows the graph of new cases over time all over the world and the present condition prevails in the world under COVID 19 as it shows the number of increasing cases day by day. Table. 1 shows the top 10 countries highly affected by COVID 19 as well as of now overall an approximate total of 2,69,564 people died due to the corona virus as of now (8th May, 2020) as indicated by the official websites.

As of now in United States more number of people died due to the virus because of certain reasons like no lockdown and no social

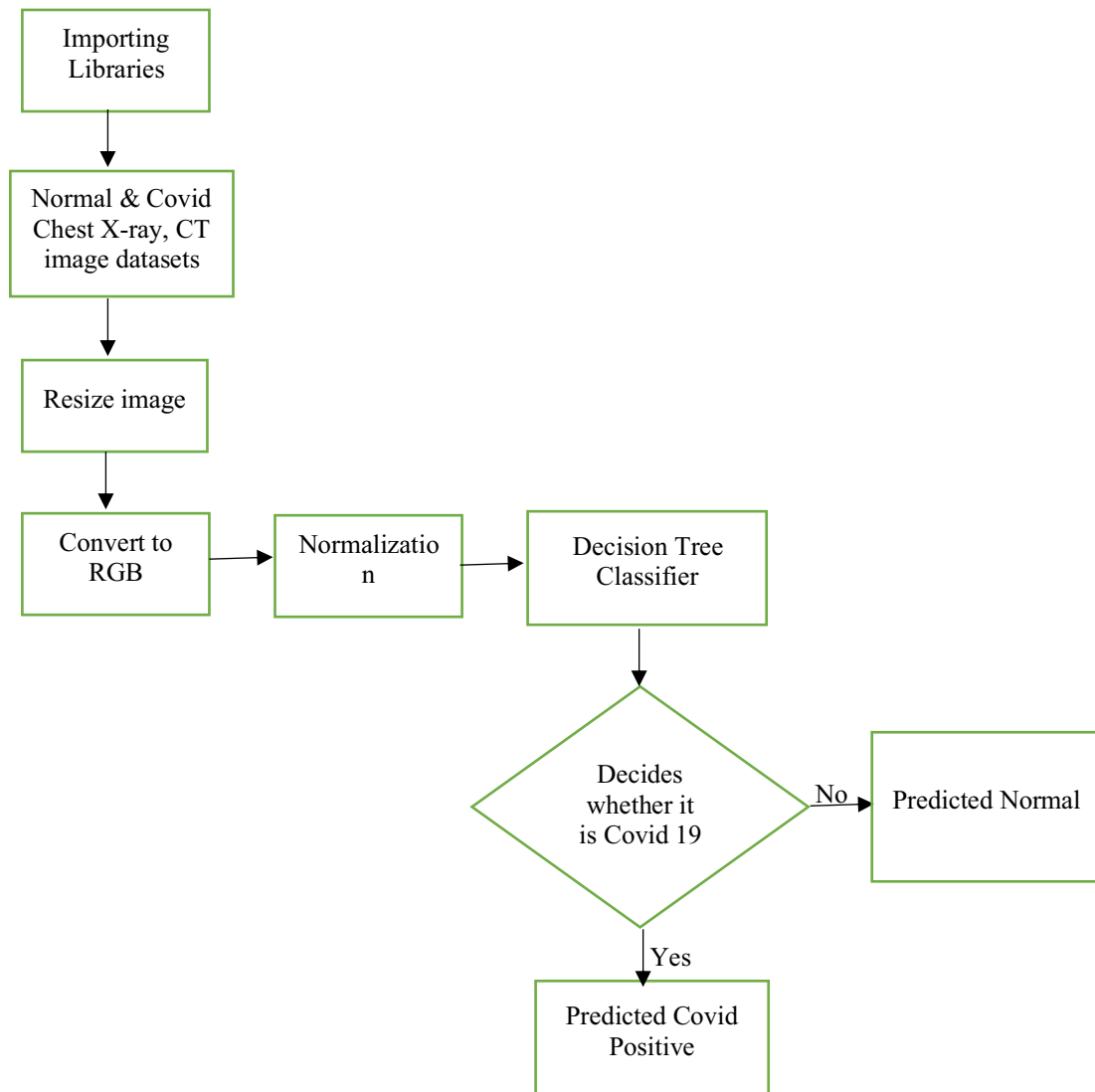


Fig. 4. Flow chart of proposed model.

Table 1

Shows top 10 countries highly affected by Covid-19 as of 8th May, 2020 [5].

Location	Confirmed	Cases per 1 million people	Recovered	Deaths
Worldwide	38,45,607	495	12,82,930	2,69,564
United States	12,90,222	3,915	1,74,709	76,600
Spain	2,22,857	4,732	1,31,148	26,299
Italy	2,15,858	3,583	96,276	29,958
United Kingdom	2,06,715	3,112	—	30,615
Russia	1,87,859	1,280	26,608	1,723
Germany	1,69,555	2,039	1,35,918	7,392
France	1,37,779	2,054	55,027	25,897
Brazil	1,36,519	646	55,350	9,265
Turkey	1,33,721	1,608	82,984	3,641
Iran	1,04,691	1,256	83,837	6,541

Table 2

Shows qualitative analysis of the proposed model with AI.

Parameters	X-ray normal / X-ray Covid-19	CT Covid-19 / CT non Covid-19
Test score	0.87	0.82
Train score	1.0	1.0
Precision score	0.88	0.77
Recall score	0.83	0.93
F1 score	0.85	0.74

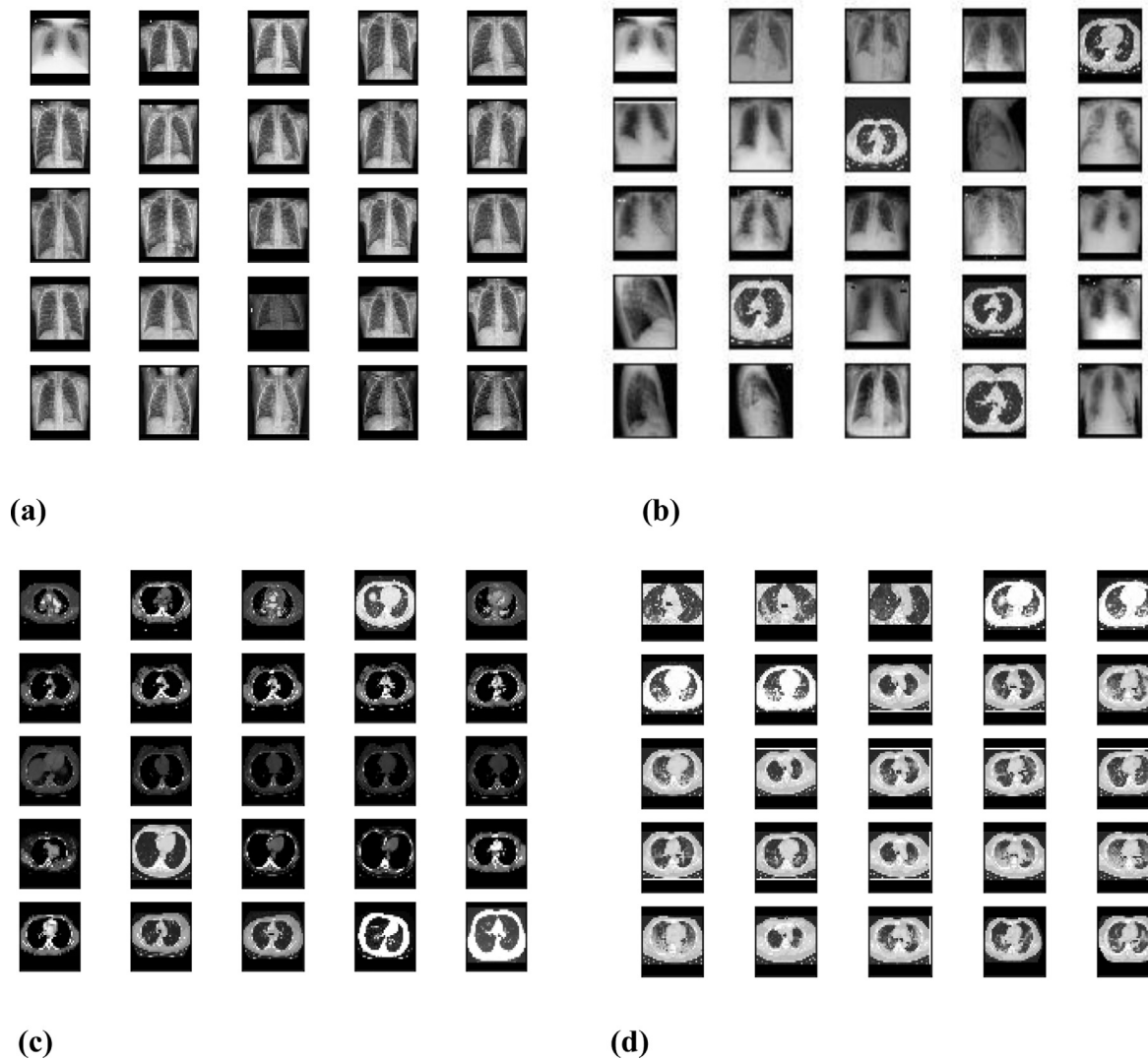


Fig. 5. (a) Normal Chest X-ray reports, (b) Covid-19 Positive chest X-ray reports, (c) CT scan non-covid-19 reports, (d) CT scan Covid-19 positive reports.

distance. These two things play a critical role to avoid spreading of virus. In some countries like India, government instituted quick actions and announced lockdown to stop spreading of virus. Owing to such swift actions in India death ratio is plausibly less although the number infected cases are rapidly increasing at the moment with a current total of 15,83,792 confirmed cases [6].

2. Methodology

In this paper, a classification method is proposed to predict Covid-19 positive infected patients by using chest x-ray and CT scan images. At initial stage the dataset images are resized as an initial requirement for the faster algorithmic thinking and subsequently resized images are converted into RGB format. Due to the prediction of the images, datasets are split into both training and testing. Final step to apply decision tree classifier along with AI to predict whether the image is normal or COVID 19 positive. Fig. 3 shows the block diagram of the recommended approach.

In this paper, a framework dependent on deep learning is produced for the recognizable of COVID 19 as a characterization task. In this examination, we arranged two sets of dataset of Chest X-ray images. The first dataset contains 9 number of Covid-19 positive and 9 number normal X-ray chest images. These datasets are gathered from Kaggle repository. The normal dataset contains 104

images and Covid-19 chest X-ray image dataset contains 297 images. The two datasets are analyzed independently in the proposed models. We utilize this dataset for profound component extraction dependent on deep learning strategy. Fig. 4 illustrates the whole process flow of the proposed model.

Deep learning element extraction depends on the extraction of highlights procured from a pre-trained CNN. The deep features highlights are extricated from completely associated layer and feed to the classifier for training and testing. The profound highlights obtained from each CNN systems are utilized by Decision tree classifier. From that point onward, the characterization is performed, and the features of all arrangement models are estimated. The proposed model applies for both Chest x-ray images as well as CT scan images. The CT scan image dataset contains 350 images and non Covid-19 CT dataset contains 395 images. The below figures show the normal chest x-ray images and Covid-19 positive chest x-ray images as well as Covid-19 CT scan and non Covid-19 CT scan images (Fig. 5).

3. Results & discussion

In this investigation, we inspected the presentation of arrangement models for recognizable Covid-19 positive with the help of AI

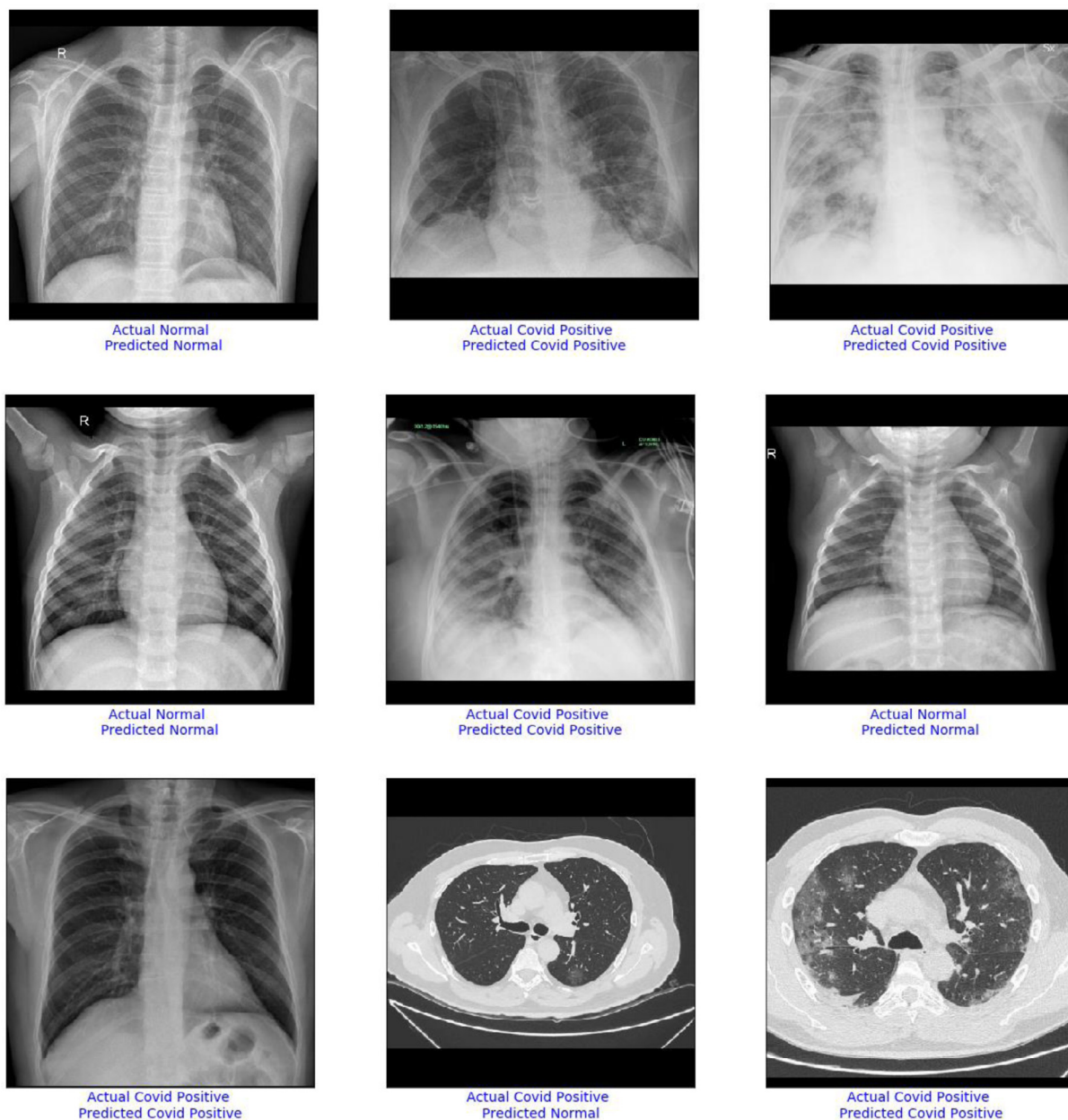


Fig. 6. Chest X-ray images predicted Covid-19 Positive with AI.

models. The test considered were executed utilizing the Anaconda-Jupyter notebook tool; all applications were run on a PC.

The estimation of execution with the help of classifier is estimated and image is analyzed in terms of F1 score, precision, and recall score. And the dataset of the images are trained and tested with the help of decision tree classifier combined with AI predicted the image whether it is normal or Covid-19 positive with precision, recall and F1 score. And the results of the output of the images are analyzed as shown in above Table 2.

Fig. 6 shows the predicted output images of the Covid-19 positive and normal chest x-ray image. Here we performed with nine chest x-ray images out of which five are correctly predicted Covid-19 positive and three images are predicted normal and one is predicted wrong. Fig. 7 shows the predicted output images of the CT Covid-19 positive and non-CT Covid-19 images as well. Here, we performed with nine are CT scan images out of which four are correctly predicted Covid-19 positive and three are predicted normal

images and two are predicted wrong. With the help of AI and deep learning strategy predicted both chest X-ray and CT scan images. The outcome depends on the information accessible in the Kaggle and Open-I according to their approved X-ray pictures and CT scan images.

4. Conclusion

Given the importance of plausible test methodology to diagnose COVID19 patients, it is imperative to identify quick testing protocol to diagnose the affected patients. The substance of the proposed model concerning the Covid-19 depends on the information accessible in WHO, the chest X-ray pictures and CT scan images are utilized for simulation intentions are gathered from Kaggle repository. For the discovery of novel coronavirus utilizing X-ray pictures and CT scan images dependent on the fusion of deep learning strategies and AI. For the decision tree classifier, we classify the feature extractions of the images and splitting for training and testing. The

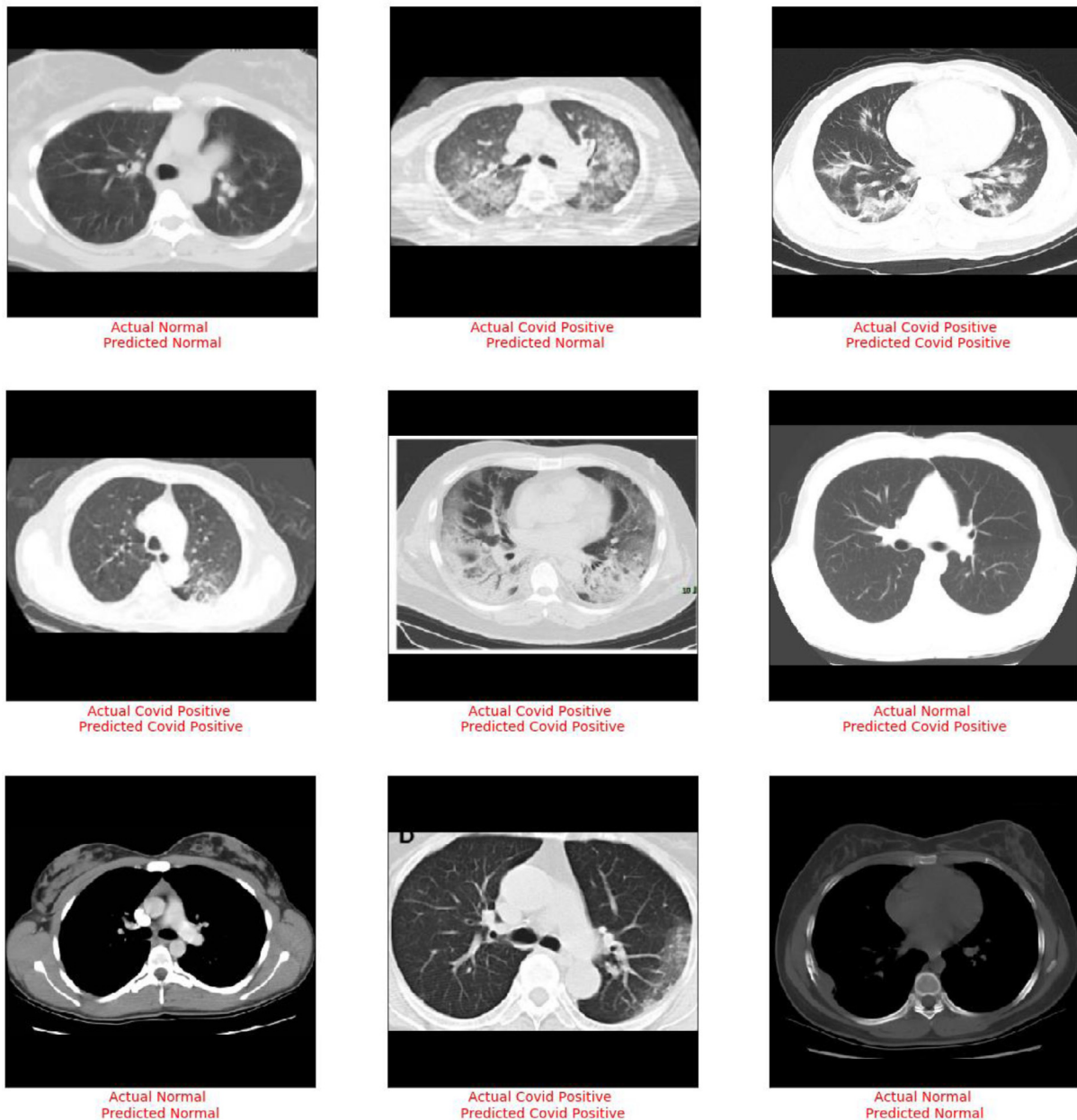


Fig. 7. CT scan images predicted Covid-19 Positive with AI.

recommended analysis model for identification of Covid-19 is accomplished to the accuracy of 93% recall score in CT scan images while 88% of precision score is verified in chest x-ray images. We believe that the present findings and its verified methodology help radiologists fast diagnosis of COVID19 infection cases. Further work is in progress involving app features of the above diagnosis to help doctors obtain predictive analytics data via internet of things (IoT).

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.

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest;

and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Acknowledgements

The authors are indebted to SRM University for supporting the work during covi-19 lockdown. One of us (DNV) is grateful to SRM University for the award of Research assistantship.

References

- [1] Sethy, PK, Santi KB. "Detection of coronavirus disease (covid-19) based on deep features." *Preprints* 2020030300 (2020): 2020. DOI: [10.20944/preprints202003.0300.v1](https://doi.org/10.20944/preprints202003.0300.v1).
- [2] Hamzah B, A F, Lau CH, Nazri H, Ligot DC, Lee G, Tan CL, et al. CoronaTracker: world-wide Covid-19 outbreak data analysis and prediction. *Bull World Health Organ* 2020(March) Submitted.

- [3] Li L, Yang Z, Dang Z, Meng C, Huang J, Meng H, Wang D, Chen G, Zhang J, Peng H, Shao Y. Propagation analysis and prediction of the Covid-19. *Infect Dis Model* 2020;5:282–92. doi:[10.1016/j.idm.2020.03.002](https://doi.org/10.1016/j.idm.2020.03.002).
- [4] Zhang X, Ma R, Wang L. Predicting turning point, duration and attack rate of Covid-19 outbreaks in major Western countries. *Chaos Solitons Fractals* 2020;135. doi:[10.1016/j.chaos.2020.109829](https://doi.org/10.1016/j.chaos.2020.109829).
- [5] Google News, Covid-19 cases situation reports <https://news.google.com/covid19/map?hl=en-IN&gl=IN&ceid=IN:en> (Accessed May 8, 2020)
- [6] Singhal T. A review of coronavirus disease-2019 (COVID-19). *Indian J. Pediatrics* 2020;87(4):281–6.