Journal Pre-proofs
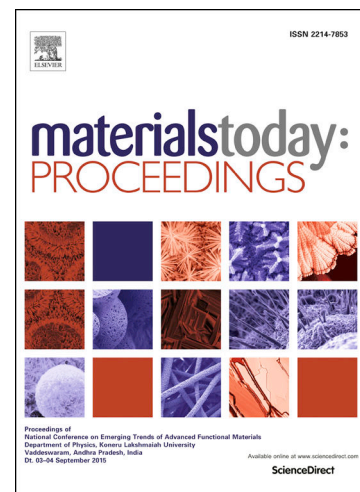
Review on Machine and Deep Learning Models for the Detection and Prediction of Coronavirus

Ahmad Waleed Salehi, Preety Baglat, Gaurav Gupta

Please cite this article as: A. Waleed Salehi, P. Baglat, G. Gupta, Review on Machine and Deep Learning Models for the Detection and Prediction of Coronavirus, *Materials Today: Proceedings* (2020), doi: https://doi.org/10.1016/j.matpr.2020.06.245

Missing

Review on Machine and Deep Learning Models for the Detection and Prediction of Coronavirus

Ahmad Waleed Salehi[a], Preety Baglat[a], Gaurav Gupta[*1]

[a, *]*Faculty of Engineering and Technology, Shoolini University, Solan, Himachal Pradesh 173229, India*

Abstract

The novel Coronavirus disease has increased rapidly in the Wuhan city of China in December 2019. This fatal virus has spread across the whole world like a fire in different stages and affecting millions of population and thousands of deaths worldwide. Therefore, it is essential to classify the infected people, so that they can take the precaution in the earlier stages. Also, due to the increasing cases spread of Coronavirus, there are only limited numbers of polymerase change reaction kits available in the hospitals for testing Coronavirus patients. That why it is extremely important to develop artificial intelligence-based automatic diagnostic tools to classify the Coronavirus outbreak. The objective of this paper is to know the novel disease epidemiology, major prevention from spreading of Coronavirus Severe Acute Respiratory Syndrome, and to assess the machine and deep learning-based architectures performance that is proposed in the present year for classification of Coronavirus images such as, X-Ray and computed tomography. Specifically, advanced deep learning-based algorithms known as the Convolutional neural network, which plays a great effect on extracting highly essential features, mostly in terms of medical images. This technique, with using CT and X-Ray image scans, has been adopted in most of the recently published articles on the Coronavirus with remarkable results. Furthermore, according to this paper, this can be noted and said that deep learning technology has potential clinical applications.

Keywords: Virus, Pandemic Disease, Machine Learning, Deep Learning, CNN, SVM;
Abbreviations: WHO, PCR, SARS, CT, AI;

1. Introduction

Viruses are a submicroscopic agent, which are made up of genetic substances inside of a protein coating and it can be found anywhere such as air, water, and soil. A virus can cause different infectious types of diseases such as common flu, cold, and warts. They also cause severe illnesses such as Acquired Immunodeficiency Syndrome (AIDS), Ebola, Smallpox [1], etc. Viruses are akin to hijackers; they enter normal cells and from those cells to replicate themselves. This process destroys or damages the cells and makes us sick, and this kind of virus is called "Pandemic virus". Various viruses attack and harm certain cells of the human body such as our respiratory system, liver, or blood. When we get a virus, it will not always affect us to get ill from the virus. the immune system of us may be capable to fight with it. For most epidemiologic infections, the only method to help is the treatment of symptoms as you should wait for the immune system to confront against or fight off the virus. For viral infections, the antibiotics cannot work, there are antiviral drugs and medicines to cure or heal some of the viral infections. The vaccines help to prevent us from getting various viral diseases.

A pandemic is an occurrence of a disease or wide-scale outbreak of any infectious disease, which can rapidly increase mortality and morbidity rate on a large-scale, over a large area. Even crossing the international boundaries, which can disturb a large population of the world and that can cause a significant social, economic, and political disturbance. Previous studies indicate that the chances of pandemics have risen in the past centuries [2]. This pandemic outbreak can happen when a novel virus turns out to be capable of circulating or spreading rapidly globally, wherein an epidemic can be in a specific region or city. In one and other cases, there is a serious illness that can rapidly spread out from one to another individual quickly. Evidence shows that the death rate of the epidemic is generally lesser than

* Corresponding author. Tel.: +91-98174-52559.
E-mail address: solan.gaurav@gmail.com

a pandemic outbreak. For example, the unfavorable pandemic in the history that caused the death of more than a hundred million people, was the Spanish flu [3].

Coronavirus has rapidly spread to almost all countries of the world to date. Coronavirus is a big family of viruses, which causes illness or sickness starting from the very common cold to very more acute or severe disease. SARS-CoV-2 showed about 80% identity of SARS-CoV-1 and 50% to the Middle East Respiratory Syndrome (MARS-CoV) from the genetic sequence and both have their origin from bats. According to the phylogenetic reports and genetic sequence, Coronavirus is adequately similar to SARS- CoV. Most likely this disease has spread from bats, as proof supports that this virus has a high grade of homology of ACE2 (Angiotensin Converting Enzyme 2) receptor from a variety of animal species [4]. The common sign of the Coronavirus infection consists of high and consistent fever, Persistent dry cough, respiratory syndromes such as breathing difficulties, and breath shortness [5]. The Coronavirus infection has spread like a fire and transitioned into a globally pandemic in which, as if now, know medical researchers found any therapeutic vaccine or drug. Because of this serious issue, it is very significant to identify the diseases at its early stage, and quickly isolate people who are infected from the non-affected population.

The Coronavirus should be the starting of an exciting time or decade in science and medicine because with the development and improvement of numerous digital technologies they can be applied and used to tackle different diseases and major clinical problems. These are the technologies that include Big-data analytics, IoT (Internet of things), Blockchain, AI which uses machine learning (ML) and deep learning (DL), etc. These are the advanced technologies that are being used in many aspects of healthcare, especially AI and machine learning systems to predict the outcomes and understanding healthcare trends [6] and how AI-based technology and deep learning can help and enhance in the diagnosis of this Coronavirus? By using a different kind of radiographical images. CT and X-ray scans of the chest can be used in diagnosing the pneumonia. There are AI-based automated CT or X-ray images analysis techniques for the detecting, monitoring, and quantifying of Coronavirus, and to find out the patients from the healthy person have developed [7].

1.1. Coronavirus Outbreak: Its Epidemiology and Pathogenesis

As of 14 April 2020, the total number of confirmed cases is 1 776 867, the confirmed number of deaths is 111,828 and the overall number of countries affected by this virus is 213 as reported by WHO. The USA, China, Italy, Spain are the most affected countries. Initially, in European countries, the very fast and worst affected country was Italy, where the public health departments, hospitals, emergency medical systems are struggling to deal with the surge of affected patients. The fatality rate of the Coronavirus was increasing in the outbreak region of China. The first 17 deaths reported by the National Health Commission of China was on 22nd January 2020 and 56 deaths on 25th January 2020 [4]. From the initial case in Dec 2019 to the rise of new incidents outside the city of Wuhan by 13th January 2020, 41 number of cases or incidents were confirmed. The epidemiologic analysis has shown which already in this early stage, human-to-human space and transmission have been happened by close interaction. On 13 January 2020, in Thailand, the first case which was outside China had been reported. This was caused by ax resident of Wuhan who traveled to Thailand, and 19 cases were reported in January from outside Wuhan City, in Beijing City [8].

Furthermore, 85.8% of 37 269 confirmed cases had either lived in or traveled to Wuhan or had close contact with persons who had been to Wuhan. The Figure 1 depicted above shows the outbreak of the disease till 9th April [8]. The Coronavirus varies widely, from the range of infection asymptomatic to severe condition and serious pneumonia along with the high number of fatality rates. The main Disease Control Centers of the Chinese reported that the majority of the number of infected people were classified as a symptom of mild to moderate condition, 13.8% as severe condition, and only 4.7% classified as critical condition. The total rate of mortality for Coronavirus confirmed cases were detected higher in males as compared to the female patients and high risk of death in both cases with rising age. The maximum fatality rate of the aged started from the 80 and above [10].
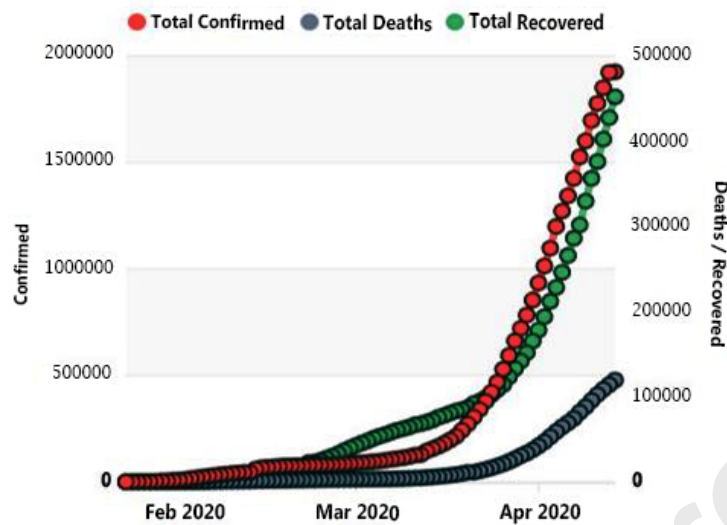
Figure. 1 Coronavirus outbreak over the time [9]

1.2. Major Prevention for Controlling Sars-Cov-2

It was very important to control the epidemic and lift the quarantine, but only if specific measures could have been taken as much as possible usage of facial masks and restricting public contact is the most effected precaution that people can take until the vaccine is not available. Apart from these, WHO announced other very significant measures to be taken by everyone such as frequent hand washing, use of disinfectants, avoid from touching the eye, face, nose and mouth, use gloves to name a few which are adopted worldwide to reduce the probability of 2nd wave of this pandemic.

In the recent report of WHO, 78% of infected cases were because of household spread or transmission and in the recent clinical report, 5% of hospital-acquired infections from patients and medical staff. Until the vaccine is made available, the studies recommended that keeping a reduction in person to person contact (less than the pre-quarantine level) is very important to contain the spread of the virus.

The study has shown that from the previous experience, Coronavirus symptoms are mainly highlighted by checking the travel history rather than radiography of chest. The treatment and detection for the isolation of Coronavirus patient cases at an early stage are very important. All hard works are being made to decrease the fast spreading of the disease and to provide time for better preparing of healthcare systems and also to the general public. It is important to clearly characterize Coronavirus to guide and help general public health recommendations and to build vaccines, therapeutics, and timely diagnostics.

Lastly, with the improvement and advancement of internet communication, it increases the availability and broadcasting of knowledge. The Internet has the possibilities in terms of development and the spreading of fake news or misinformation. So, the government sector must be responsible for providing and delivering the accurate information and clarifying misinformation to help the general public to face this Coronavirus infection. The below figure depicts the percentage of the Coronavirus symptoms.
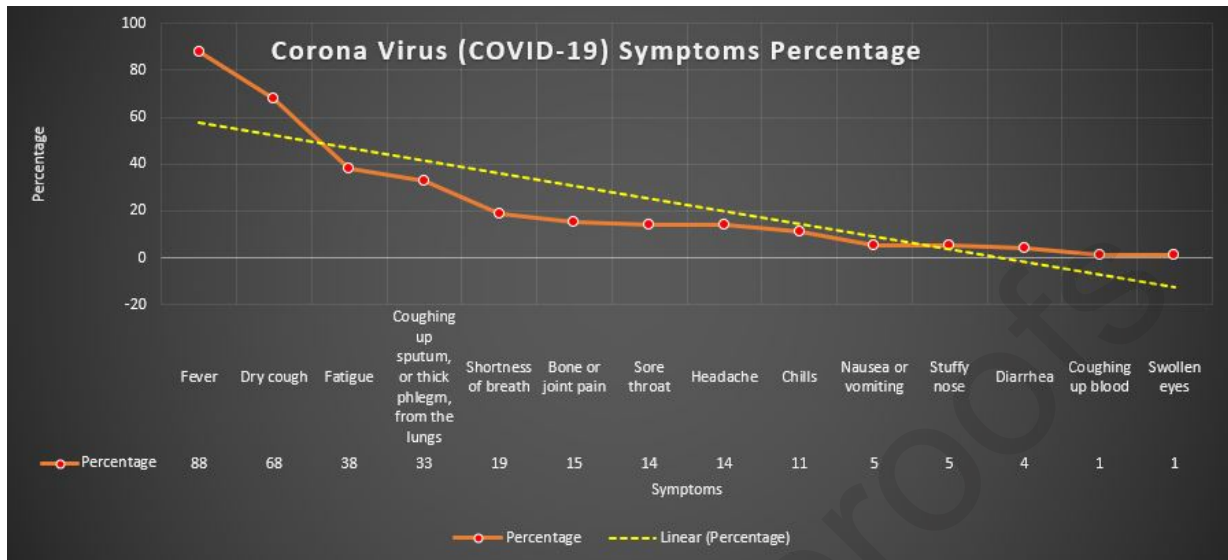
Figure. 2 Coronavirus symptoms percentage

2. Literature Review

In this article, we have selected publications for review of literature that have assessed machine and deep learning algorithms applied in medical images to solve a clinical problem and compared each algorithm regarding their output and performance.

Machine learning is the branch of AI which is based on how the system can learn from previous data, recognize patterns, and to make decisions with minimum human intervention, for example, algorithms includes Support Vector Machine (SVM), logistic regression, clustering, etc. Deep learning (DL) is a subset of machine learning, and in terms of medical images, it can be defined as a computational model [20, 21] which is composed of several processing layers to learn data representation and extract the features with multiple abstraction levels. An example of a deep learning type is CNN [11].

In [12] has proposed a deep learning-based technique called deep transfer learning, which can predict patients with Coronavirus disease automatically. It uses images of chest X-ray obtained from patients with Coronavirus and from a healthy person. Dataset of 50 patients, with Coronavirus, images of X-Ray were taken from a shared GitHub repository and 50 X-ray images of healthy humans have taken from a repository in Kaggle. In this study, results showed that the pre-trained model, ResNet50, yielded 98% accuracy among the other three models. They have also stated that we believe in our findings and this can help the public health care workers to make decision in clinical practice because of its high performance and accuracy. Also, for patients with Coronavirus, early prediction of the infection can avoid the rapid spreading of the disease.

This paper study [13] presented automated techniques are used for classifying the X-Ray of chest into pneumonia and the class of disease-free by using nine architectures of deep learning, which includes the following,

- Baseline CNN
- DenseNet201
- VGG16, VGG19
- Inception_ResNet_V2, Inception_V3
- Xception
- Resnet50
- MobileNet_V2

Experiments have been conducted using a CT scan and X-Ray dataset that includes 5856 images with 1583 normal and 4273 pneumonia, and the performance was evaluated using different performance metrics. The result shows that the MobileNet_V2, Inception_Resnet_V2, and Resnet50 have given more accurate results, which is more than 96%. In addition, they have suggested that with the bigger datasets and more sophisticated techniques of feature extraction based upon deep learning image segmentation may improve the performance.

In [14] established a deep learning paradigm for the screening of Coronavirus patients at an early stage. The main aim of this paper is to distinguish Coronavirus from Influenza-A viral pneumonia and normal cases with the use of CT images. CT samples have taken from three hospitals designated to Coronavirus from China, Zhejiang Province. The total number of 618 samples was collected, which includes 219 from 110 Coronavirus patients, 224 samples of CT from patients with the viral pneumonia Influenza-A and CT samples of 175 healthy people. The experiments result of this study shown an overall 86.7% accuracy from the viewpoint of CT cases as a whole. They demonstrated that it can be a promising accompanying diagnostic tool for the clinical frontline doctors.

As in [15] introduced a new approach of deep learning called COVIDX-Net to help radiologists to automatically diagnose patients with Coronavirus using X-ray images. This technique is built on seven Deep CNN classifiers, which consists of DenseNet121, VGG19, ResNetV2, InceptionResNetV2, Xception, InceptionV3 and MobileNetV2. In this model, they have successfully experimented and evaluated the result which is based on 80% training phase, 20% testing phase of X-ray images. The VGG19 and the Dense CNN (DenseNet) models have shown a better and similar functioning of automated classification of Coronavirus with 0.89 f1-scores for normal and 0.91 for Coronavirus, respectively. X-ray images of the public dataset used in this work, which consists of 50 images, divided into two classes as normal cases of 25 and positive Coronavirus images of 25.

Another paper implements an automated detection system of Coronavirus diseases using Deep learning Convolutional neural network and chest X-ray images. In this paper, they have implemented three different CNN based models such as ResNet50, Inception- ResNetV2, and InceptionV3 for the identification of Coronavirus cases using X-ray images. The dataset of 50 Coronavirus patients has been taken from the open-source GitHub repository (Dr. Joseph Cohen) and 50 images of normal chest X-rays have been selected from a repository in Kaggle (Chest X-Ray images). The overall performance out of three models shows that the best results obtained from the pre-trained model of ResNet50 with an accuracy rate of 98% and the other two models accuracy rate from Inception-ResNetV2 and InceptionV3 are 87% and 97%. In future work, they have mentioned that the different CNN models classification performance can be tested by adding the greater number of images in the dataset [7].

In this study [16] they present earlier detection of Coronavirus using CT Scan images by using machine learning methods and the dataset is taken from "Societa-Italiana di Radiologia Medica-e Interventistica", which belongs to the 53 Coronavirus cases and data involving CT abdominal images of 150 and 150 CT images. Then, patch regions of the images were cropped, and four different subsets of the patch were created. The feature extraction techniques used in this study are the following,

- Grey -Level Co-Occurrence Matrix,
- Local Directional Pattern
- Grey-Level Size-Zone Matrix
- Grey-Level Run-Length Matrix
- Last-Discrete Wavelet Transform

The SVM classifier was applied to classify the extracted features and the best classification outcomes were obtained from GLSZM future extraction techniques with an accuracy rate of 99.68%. In future work more classification and segmentation, studies should be done on Coronavirus disease by increasing the dataset.

As in [17] used different deep learning-based techniques for Coronavirus detection diseases by using X-ray images. Different classification models, based on SVM, using deep features of various deep learning architectures as follow,

- AlexNet
- VGG16, VGG19
- GoogleNet
- ResNet18, ResNet50, ResNet101
- InceptionV3
- Inception ResNetV2
- DensNet201

- XceptionNet

are used for detecting the Coronavirus patients. For experimental work they have been conducting, they use an X-ray images dataset that includes 50 images with 25 Coronavirus cases from GitHub (Dr. Joseph Cohen) and 25 normal cases from Kaggle (X-ray Images of Pneumonia). The results show that SVM + ResNet50 obtained high accuracy (FPR=95.52%, F1 score= 95.52%, MCC = 91.41% and Kappa= 90.76%) for detecting the Coronavirus patients as compare to the other models.

3. Result Analysis

Various ML and DL-based algorithms and techniques used for classification of the novel disease known as Coronavirus 2019 have been studied and reviewed. Different papers are studied in which the majority of the related paper used different deep learning architecture. According to the literature review, it is demonstrated that deep learning with convolutional neural networks might have remarkable impacts on the automatic detection and automatic extraction of highly essential features from chest images which is related to the diagnosing of Coronavirus. The below figures 3, 4 demonstrated the accuracy of different models using CT, X-Ray images used for the detection of this virus.
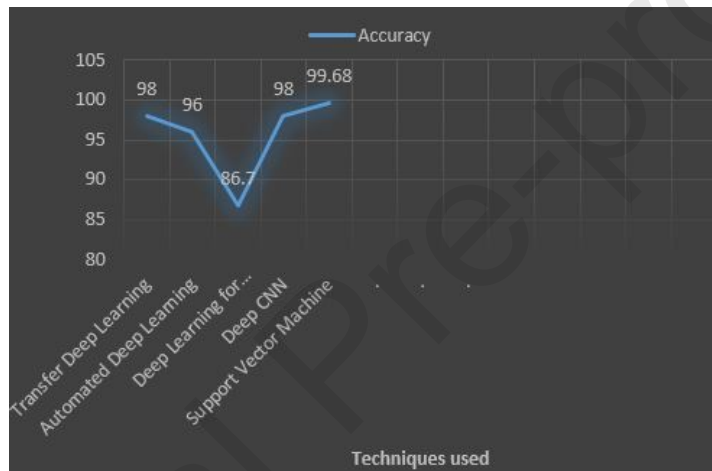

Figure. 3 Comparison between techniques and accuracy.

Figure. 4 Comparison between data type and accuracy.

As in the following table on the next page, we have categorized each paper with their what technique, data modality, data source they have used, and the result they obtained, respectively with the given references.

Table 1. Comparison of various techniques used for COVID-19 detection

Ref	Technique	Data Type	Data Source	Accuracy
[12]	Transfer Deep Learning for automatically predicting COVID-19	X-Ray	Kaggle and GitHub	98%
[13]	Automated Technique for Detecting and Classifying Pneumonia-based using Deep Learning	CT and X-Ray	X-Ray, CT Dataset publicly available on the internet	96%
[14]	Deep Learning for Screening COVID-19 pneumonia	CT	Hospital of Zhejiang, China	86.7%
[15]	Deep CNN	X-Ray	X-ray images of a public dataset	VGG19, DenseNet models: f-scores=0.89 normal & Coronavirus-19= 0.91
[7]	Automated Deep Convolutional Neural Network	X-Ray	50 Coronavirus patients (GitHub) 50 normal X-ray (Kaggle)	98%
[16]	Support Vector Machine	CT	Total = 150 CT images Coronavirus = 53	Classification accuracy result obtained from GLSZM=99.68%
[17]	Support Vector Machine based on deep learning approach (Deep Features)	X-Ray	Coronavirus cases = 25 Normal cases = 25 (GitHub, Kaggle)	Accuracy: SVM+ ResNet50 (FPR=95.52%, F1 score=95.52%, MCC = 91.41% and Kappa= 90.76%)

4. Conclusions and Discussion

For the treatment of Coronavirus patients, there is no effective drug present until now and due to the rapid increase in the number of cases of Coronavirus patient's effectual medicinal approach is urgently needed to treat the patients worldwide [18]. Earlier detection and prediction of Coronavirus cases can decrease the spread of diseases [7]. Worldwide patients of Coronavirus data will be a great source and useful for the researchers working on AI and ML to develop an automatic diagnostic tool, therapeutic strategy against Coronavirus patients, and for the same type of pandemics for the coming future [18, 19].

In this study, we have summarized different AI-based approaches and the current situation of spreading Coronavirus. According to the results and studies reviewed, it is indicated that using convolutional neural networks,

which is the deep learning-based technique, might have significant effects on automatic tools in terms of detection, distinguishing, and extraction of essential features automatically from the X-Ray images that are related to Coronavirus diagnosis.

Due to the early stage of Coronavirus, there are still some limitations of the related studies that can be conquered or overcome in the future researches. Specifically, an in-depth analysis requires a lot more patient data, particularly those patients suffering from the novel Coronavirus. So, the main point of focus for research work in the future could be distinguishing the patients indicating mild symptoms, instead of pneumonia symptoms, whereas these symptoms might not be visualized on X-rays more accurately, or even not at all to be visualized. As more authentic dataset will be available in the near future, a more accurate AI-based prediction model can be formed. We are working in this direction to procure a more authentic dataset as on today after reviewing all the papers mentioned in Table 1, we are working to develop CNN based deep learning methods for the prediction of Coronavirus patients.

AI-based approaches are very useful for the automatic detection of Coronavirus patients using X-ray and CT images. The important point is to increase the number of datasets for Coronavirus patients and using advance deep learning algorithms to achieve better performance for the detection and prediction of Coronavirus. Also, in spite of the fact that proper treatment or cure cannot be determined only from an X-Ray or CT image, these techniques would be useful as an initial screening of the patients. Appropriate use of quarantine actions should be implemented on the positive samples, until a specific treatment and a complete examination procedure is followed.

Acknowledgments

References

[1] D. J. Wilson, D. Falush, and G. McVean, "Germs, genomes and genealogies," Trends Ecol. Evol., vol. 20, no. 1, pp. 39–45, 2005.

[2] T. M. Tumpey, A. Garcia-Sastre, J. K. Taubenberger, P. Palese, D. E. Swayne, M. J. Pantin-Jackwood, S. Schultz-Cherry, A. Solorzano, N. Van Rooijen, J. M. Katz, and C. F. Basler, "Pathogenicity of Influenza Viruses with Genes from the 1918 Pandemic Virus: Functional Roles of Alveolar Macrophages and Neutrophils in Limiting Virus Replication and Mortality in Mice," J. Virol., vol. 79, no. 23, pp. 14933–14944, 2005.

[3] A. Erkoreka, "The Spanish influenza pandemic in occidental Europe (1918-1920) and victim age," Influenza Other Respi. Viruses, vol. 4, no. 2, pp. 81–89, 2010.

[4] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," J. Autoimmun., vol. 109, no. February, p. 102433, 2020.

[5] C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges," Int. J. Antimicrob. Agents, vol. 55, no. 3, p. 105924, 2020.

[6] Z. Wang and K. Tang, "Combating COVID-19: health equity matters," Nat. Med., vol. 26, no. April, pp. 2019–2021, 2020.

[7] "Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Turkey."

[8] J. Sun, W. T. He, L. Wang, A. Lai, X. Ji, X. Zhai, G. Li, M. A. Suchard, J. Tian, J. Zhou, M. Veit, and S. Su, "COVID-19: Epidemiology, Evolution, and Cross-Disciplinary Perspectives," Trends Mol. Med., pp. 1–13, 2020.

[9] "COVID-19 related analytics, graphs, and charts   Corona Tracker." .

[10] "Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study," Bmj, vol. 1295, no. March, p. m1295, 2020.

[11] M. Nagendran, Y. Chen, C. A. Lovejoy, A. C. Gordon, M. Komorowski, H. Harvey, E. J. Topol, J. P. A. Ioannidis, G. S. Collins, and M. Maruthappu, "Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies in medical imaging," BMJ, vol. 368, pp. 1–12, 2020.

[12] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," Phys. Eng. Sci. Med., no. 0123456789, pp. 1–6, 2020.

[13] K. El Asnaoui, Y. Chawki, and A. Idri, "Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning," 2020.

[14] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, K. Xu, L. Ruan, and

W. Wu, "Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia," pp. 1–29, 2020.

[15] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," 2020.

[16] M. Barstugan, U. Ozkaya, and S. Ozturk, "Coronavirus (COVID-19) Classification using CT Images by Machine Learning Methods," no. 5, pp. 1–10, 2020.

[17] P. Kumar and S. Kumari, "Detection of coronavirus Disease (COVID-19) based on Deep Features," Https://Www.Preprints.Org/Manuscript/202003.0300/V1, no. March, p. 9, 2020.

[18] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng, "Artificial intelligence and machine learning to fight COVID-19," Physiol. Genomics, vol. 52, no. 4, pp. 200–202, 2020.

[19] A. W. Salehi, P. Baglat, and G. Gupta, "Alzheimer's Disease Diagnosis using Deep Learning Techniques," Int. J. Eng. Adv. Technol., vol. 9, no. 3, pp. 874–880, 2020.

[20] Gupta, Gaurav, Ankit Gupta, Varun Jaiswal, and Mohd Dilshad Ansari. "A Review and Analysis of Mobile Health Applications for Alzheimer Patients and Caregivers." In 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 171-175. IEEE, 2018.

[21] Sharma, Lalaantika, Gaurav Gupta, and Varun Jaiswal. "Classification and development of tool for heart diseases (MRI images) using machine learning." In 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 219-224. IEEE, 2016.

Declaration of Interest Statement

Date: 04/05/2020

To

The Editor,
Materials Today: Proceeding Journal.

Dear sir,

We are authors, submitting the manuscript for publication in your journal. This is our research manuscript and has not been published or considered for publication by any other journal or elsewhere. Kindly consider the manuscript for publication in your journal.
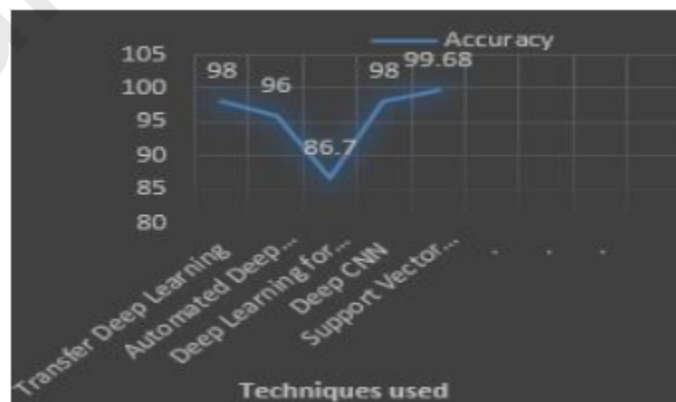
Thanking you,

Yours Faithfully,

1. Ahmad Waleed Salehi
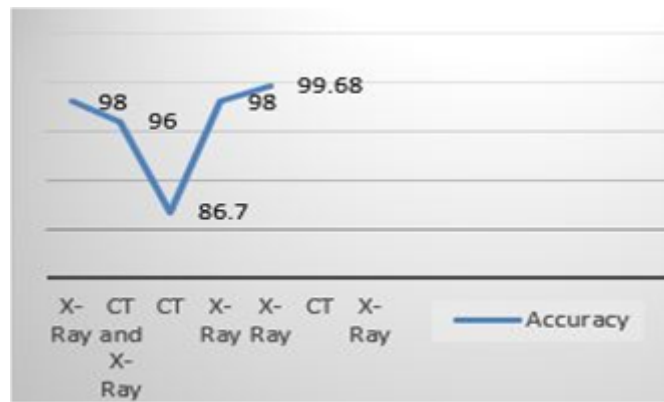2. Preety Baglat
3. Dr. Gaurav Gupta
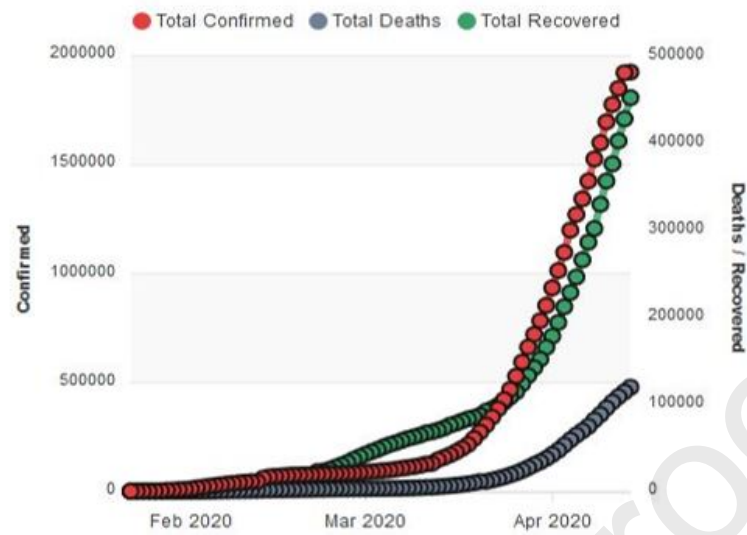
Fig. 1 *COVID-19 outbreak over the time*

Fig. 2 *Coronavirus symptoms percentage*

Fig. 3 *Comparison between techniques and accuracy*

Fig. 4 *Comparison between datatype and accuracy*