Detection and Prevention of Covid-19

Group Members : Amitesh Patil, Kshitij Patil, Pranit Patil,

Vivek Pundkar

SY, ET-C, Vishwakarma Institute Of Technology, Pune.

***Abstract*—** ***This study aims to investigate whether using machine learning and in-depth reading methods on chest X-ray images can detect coronavirus cases. Chest X-ray data sets were obtained from Kaggle and GitHub and were pre-processed into a single database using random samples. We have used a number of learning methods and in-depth learning methods that include Convolutional Neural Networks (CNN) and students of archeology. In an in-depth learning process, several pre-trained models are also used to transfer learning to this database. Our proposed CNN model showed the highest accuracy (94.03%), AUC (95.52%), f-measure (94.03%), sensitivity (94.03%) and precision (97.01%) and the lowest collapse (4.48%) and the missed rate (2.98%) respectively. We also tested specifications and fall rates and accuracy to identify non-COVID-19 people more accurately. As a result, our new models can help detect COVID-19 patients early and prevent further public outbreaks compared to conventional methods.***

***Keywords—CNN , Covid-19 ,Dataset , X-ray , Machine Learning , ResNet18***

1. INTRODUCTION (*HEADING 1*)

Novel coronavirus (COVID-19) is a progressive pandemic caused by acute acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2). It was first discovered in Wuhan, China in December 2019 and has retained phylogenetic similarity to Severe Acute Respiratory Syndrome (SARS-CoV). Coronaviruses first discovered in the 1960s in patients' nostrils. It is a large infectious group that often infects RNA viruses caused by respiratory, liver and neurologic infections among humans and other mammals. Coronaviruses belong to a large family of viruses that cause diseases such as Middle East Respiratory Syndrome (MERS-CoV) and SARS - CoV. SARS-CoV, MERS-CoV and SARS-CoV-2 come from bats. SARS-CoV-2 is the third most common form of human coronavirus in the last two decades, preceded by outbreaks of SARS-CoV and MERS-CoV in 2002 and 2012 respectively. Similarly, COVID-19 is a highly contagious infectious disease worldwide. The World Health Organization (WHO) has declared the outbreak as a global public health emergency on January 30 and as a pandemic on March 11, 2020. To control these diseases, the WHO is urging people to take precautionary measures, such as maintaining contact. and community, washing hands with soap and disinfectant, avoiding contact with noise, mouth, eye, etc. However, most affected countries are closed to their regions. preventing the spread of the disease. Patients with Coronavirus can be diagnosed mainly with common symptoms such as fever, cough, fatigue, loss of appetite, muscle pain etc. The total number of patients infected with the virus is estimated to be 4,885,738 on May 19, 2020 where the percentage of total recovery, mortality, morbidity and critically ill patients were defined as 86%, 14%, 98% and 2% respectively. To reduce the epidemic, it is necessary to identify patients with COVID-19, isolate and confirm treatment policy in the early stages. There exists two types of testing procedures namely i) Molecular

diagnostic tests, and ii) Serologic tests . Cell diagnostic tests can identify those people who have been infected during the test. Reverse Transcription Polymerase Chain Reaction (RT-PCR) is a molecular diagnostic test currently considered to be a gold standard in the diagnosis of COVID-19 that detects viral RNA from the sputum or nasopharyngeal swab. It is associated with a fair amount of good truth and requires a certain amount of equipment. Another method currently under development could detect viral proteins and detect COVID-19 called viral antigen detection. Instead, once a patient has recovered, cell tests can no longer detect that the person has been infected before. Antibody growth may be indicated by its response over a period of time and depends on the host. Serologic tests can detect antibodies that are key diagnostic tools for patients with COVID-19. However, these tests are time consuming and not available to most people, especially in low- and middle-income countries due to a lack of laboratory and human resources. However, we need to take other procedures to find patients as soon as possible.

1. LITERATURE REVIEW

Later, demonstrating the importance of X-ray images and in-depth learning in the diagnosis of COVID-19, a few visual studies will be discussed in this section. Recently, a study by Hamdan et al. introduced COVIDX-Net, which contains a comparative analysis of seven in-depth study models for COVID-19. The models used were VGG19, DenseNet201, ResNetV2, InceptionV3, InceptionResNetV2, Xception, and MobileNetV2 using a binary data set consisting of 50 X-ray samples (25 healthy and 25 COVID-19). The experiments were performed using X-ray images from two sets of data: the COVID-19 X-ray image site [30], which included 123 X-rays, and an Adrain Rosebrock data set. The study achieved 90% accuracy with VGG19 and DenseNet201. However, research suffers from limited data set limitations. Similarly, the VGG19 and ResNet50 models were used and compared with the proposed COVID-Net model in a study conducted by Wang et al. in COVID-19 diagnostics using a well-trained ImageNet model and Adam optimizer using a multiclass data set (normal, pneumonia, and COVID-19). They found a better accuracy of 93.3% compared to the study mentioned earlier. There were 13,975 X-ray images taken from most open source data sets. However, to address the problem of data inequality a data expansion approach was used. Like Wang et al., Another study by Apostolopoulos et al. also used VGG19. The data set used in the study was collected from four open source data sets, Radiopaedia, and the Italian Society of Medical and Interventional Radiology (SIRM); the total number of X-ray images was 1427 (224 COVID-19, 700 pneumonia, and 504 normal). Tests are designed for binary and multi-stage categories.The highest

accuracy achieved for the binary class was 98.75% and for multiclass, the highest accuracy was 93.48%. Like the two previously mentioned studies, VGG19 performed much better than other models with accuracy.

Undoubtedly, the importance of chest X-ray in the diagnosis of COVID-19 and the impact of the in-depth model of automatic X-ray analysis have prompted the need for further testing. Apart from these benefits, there is a paradox of finding an open source data set that contains a large number of COVID-19 X-ray images. Most previous studies suffer from a small set of data or data inequalities. To avoid these pitfalls, we have used a combination of data sets that are a combination of open source data sets. A lot of research has been done, but really, there is still a need for further research.

The analysis and discovery of COVID-19 has been extensively investigated over the past few months. The first part of this section discusses problems related to the detection of COVID-19 based on in-depth study methods using CT scans and chest X-rays. The second section reviews related literature to assess future estimates of the value of COVID-19 certification, recovery, and death. COVID-19 is now a global epidemic due to its rapid spread. It is a great challenge to find people who have been exposed because they do not show symptoms immediately. Therefore, it is necessary to find a way to estimate the number of people who may be infected regularly in order to take appropriate action. AI can be used to test a person with COVID-19 as an alternative to the more time-consuming and expensive methods. Although there are few studies on COVID-19, this study focuses on the use of AI in predicting COVID-19 events and identifying patients infected with COVID-19 through chest X-ray images. Several research sites have developed AI (e.g., diagnostic tests in health care). One of the main advantages of AI is that it can be used in a trained model to distinguish invisible images. In this study, AI was used to determine if a patient had COVID-19 using a chest X-ray image. AI can also be used to predict (e.g., how population will increase over the next 5 years) based on available evidence. Thus, predicting the near future may help authorities to take necessary action [14]. Wynants et al. focused on two key concepts. The first concept involved studies related to the discovery of COVID-19, and the second involved studies related to predicting the number of people to be infected in the coming days. Research analysis has emphasized that most existing models are accurate and biased. The authors suggested that COVID-19 research-based data should be made available to the public in order to encourage the adoption of specially designed models for acquisition and assumptions.

The main purpose of reviewing articles on AI techniques used in the acquisition and classification of COVID-19 clinical images is to understand current thinking in the field and to justify the need for future research on related issues that have been overlooked or poorly studied. This collection contained only one article. In Ref. [9], research has reviewed the rapid responses to the medical imaging community (AI-enabled) in relation to COVID-19. The authors emphasized that AI-enabled image detection could greatly assist in performing the scanning process and shaping work with minimal contact with patients, providing the best possible results.

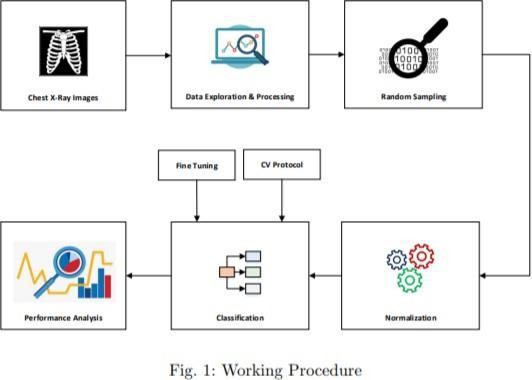
protection of professional photographers. They focus on a wide range of medical imaging and diagnostic techniques involved in COVID-19, including imaging, classification, diagnosis and follow-up, using AI and X-ray integration and CT imaging.

Recently, several studies have investigated COVID-19 clinical images using a variety of methods based on in-depth study. Wang et al. [25] introduced an in-depth open-source study model and identified a large benchmark database called COVIDx with 13,975 patient chest X-ray images to test COVID-19. This model was not only verified by the extensive details of the COVID-19 sensitive material but also extracted relevant information from the images examined. Their research produced 93.3% accuracy over the COVIDx database. Maghdid et al. [3] used a convolutional neural network (CNN) and transmitted the study using AlexNet for X-ray and CT image databases. They identified 94% (88% specificity) for CNN and 98% for AlexNet (96% specificity). Subsequently, Ali Narin et al. [21] analyzed chest X-Ray images using three CNN models such as ResNet50, InceptionV3 and InceptionResNetV2 to automatically detect COVID-19. Their study showed a very high rate of 98% accuracy using 5 times cross confirmation. Abbas et al. [1] CNN's comprehensive DeTraC model aims to provide a solution by transferring information from common object recognition to domain-specific functions. Their algorithm showed 95.55% accuracy (specified 91.87%, and 93.36% accuracy). Apostolopoulos and Mpesiana [3] initiated transmission studies using CNN into a database of small medical imaging. Their data contains 1427 X-ray images including 224 certified COVID-19 images. They showed excellent accuracy of 96.78%, sensitivity of 98.66%, and specificity of 96.46% respectively. Hemdan, Shouman and Karar [11] propose COVIDX-Net which includes seven deep CNN structures that investigated 50 chest X-ray images with 25 cases of COVID-19. Their algorithm produced 90% accuracy and 91% F-score. Subsequently, Khobahi, Agarwal and Soltanalian [13] proposed a slow-track auto-based reading method that ensures an average accuracy of 93.5%. Li et al. [15] provided a low-density neural network (DNN) based on a mobile application called COVID-MobileXpert. Vertical abstracts for chest X-ray may be used in COVID-19 tests. Therefore, both ShuffleNetV2 and MobileNetV2 have high AUROC values ​​of 94% and 94.30% respectively. Minae et al. [20] proposed a in-depth reading framework based on 5000 images called COVID-Xray-5k in which they used ResNet18. Another method is currently being developed that can detect viral proteins and detect COVID-19 called viral antigen detection. Instead, when the patient recovers, cell tests can no longer detect that the person has been infected before. Antibody growth may be indicated by its response over a period of time and depends on the host. Serologic tests can detect antibodies that are key diagnostic tools for patients with COVID-19. However, these tests are time consuming and not available to most people, especially in low- and middle-income countries due to a lack of laboratory and human resources.

.

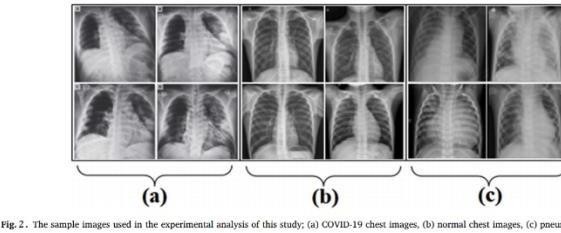
1. PROPOSED METHADOLOGY

The methodology was obtained by COVID-19 patients from publicly available databases. Figure 1 shows the working steps of our operating system. This process is divided into data collection, pre-processing, classification and sequential analysis. In this work, this approach is described in more detail in the following section:



1. *Dataset*

For this project, we have used three categories of publicly accessible databases. These classes are standard, pneumonia, and chest images of COVID-19. All data sets are X-ray images, and each image is converted to JPG format. Since COVID-19 is a relatively new disease, the number of images associated with this virus is limited. In the experimental analysis, 70% of the data was used as training data, and 30% was used as test data. Sample images of the database are shown in Fig. 2.



1. *Data Pre-Processing and Augmentation*

In In the process of pre-processing the data, we have used the familiarity of the images to generate binary values. Thereafter, the images were transformed into gray matter. However, previously trained CNN models such as VGG16, ResNet50, InceptionV3 cannot support gray images when RGB images are considered to detect COVID-19 patients. We changed the file size to 100 × 100 pixels.

Color values ​​256 (0-255) and mapped between 0 to 1 corresponding to probability theory. Therefore, it was very easy to place data at this interval. In a few cases, it improve the performance of the model depending on which opening function is used. At this stage, a data augmentation method was used, with the aim of alleviating the problem of over-modeling. Due to the deep nature of the pre-trained model, there is a high risk of overload if the data set size is small. To avoid this setback, additional images were created using augmentation of data. The data-enhancing method enhances data processing, especially X-ray data sets. Addition was done using three steps — resizing, scanning, and rotating. To change the image size, a 224 × 224 × 3 dimension was used. In addition, a horizontal rotation was used to maximize modeling in all possible areas of the COVID-19 markers on X-rays. Finally, some photographs were made using a 15-degree rotation. The additional methods used tried to improve the proposed model performance. Data augmentation is only used on X-ray training data set.

1. *Proposed Convolutional Neural Network*

Convolutional Neural Networks (CNN) is a special class of artificial intelligence networks (ANN) that can take pictures as input and investigate them (see Figure 3). In CNN design, there is little communication between layers and weights that are shared between neurons that emerge from hidden layers. Provides excellent performance in computer vision and image analysis such as image recognition, object detection, semantic classification and medical image analysis etc. It considers the input layer by maintaining a sequence of hidden and outgoing layers. Like a regular ANN, CNN contains a sequence of hidden layers and is basically described as a convolutional and polling layer. Therefore, the functioning of these layers is called convolutional and voting function respectively. Alternatively, they piled up to lead a series of fully connected layers followed by an exit layer.

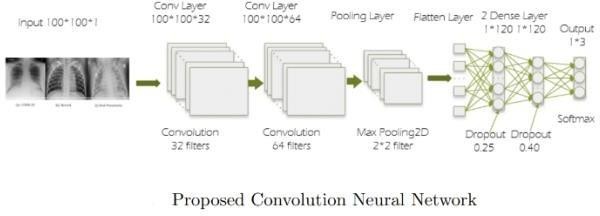
Considered two layers of convolution in our proposed CNN structure. Input images are transferred to the first layer of convolution, and the data size is 100 × 100.

× N when N was the number of hyper spectral image channels in addition to considering gray scale images as N inputs

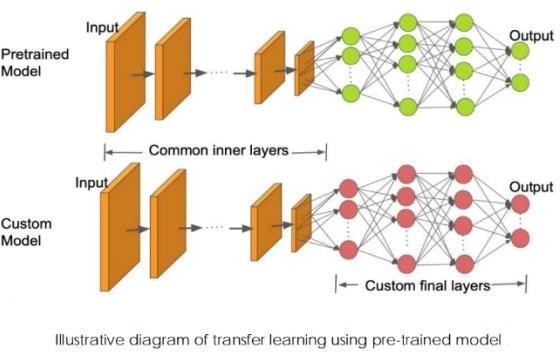
= 1. With the input value, we have created multiple channels to generate output across all layers. The detection of any hidden layer means that every outgoing neuron was connected to a small area in the input by a weight matrix called a filter or kernel. Then, describe the multiple kernels throughout the convolution layer in each output. The weight matrix or filter or kernel circled all the regions or area of ​​the image and proceeded to duplicate by the pixel values, compiled and the total output can produce a corresponding volume of output. Each filter is moved closer to the input leading to a single 2D output. The output is mounted on a pole resulting in the output volume corresponding to the filter. For convolutional layers, 3 × 3 kernels are used. In the first convolutional layer, there were 32 filters, so the output of the first layer was 100 × 100 ×

32. Next, the data was transferred to the second conversion layer and there was a kernel of size 3 × 3 and 64.

filters, which is why the output of the second conversion layer is 100 × 100 × 64. Then we used a maxpooling2D layer with a combined size of 2 × 2. A few dense layers are used by 120 and 60 neurons respectively. After each congested layer, a drop-off layer was provided to reduce overcrowding problems. Exit volume or 2D map size was determined and each piece was separated by a so-called feature or activation map. Finally, a dense output layer was activated by C neurons where C was the number of classes. In order to address the multi-level division, we used the SoftMax activation function and in combining the loss function according to the cross entropy category and Adam optimizer was considered in this model.



Model Development :



ResNet:

ResNet, also known as the deep residual network, was originally proposed in 2015 on the grounds of "short-term identifiable connections". It is also among the best models that use ImageNet. ResNet skips one or more layers and handles the issue of gradient depletion. Among the key benefits of ResNet is simplification. In addition, model accuracy can be improved with increasing model depth [60]. The ResNet model skips one, two, or more layers and is directly connected to any layer, not the nearest layer, using the indirect ReLu activation function. ResNet uses the forward and backward streaming method.

1. PERFORMANCE ANALYSIS

To test the reliability of the proposed detector based on in-depth study based on COVID-19, we adopted the same metrics as those used by Alazab et al and considered the following general metrics: accuracy, memory, and F rating. These metrics are calculated on the basis of true-positive (TP), true-negative (TN), false-positive (FP), and fall-negative (FN) results: - TP is part of COVID-19 images X-rays are well labeled as good. - FP is right

part of XV ray images of COVID-19 negative (negative) labels incorrectly labeled as positive. - TN is a component of X-ray (unhealthy) X-ray images that are properly labeled as healthy. - FN is the number of people with COVID-

19 Chest X-ray images incorrectly labeled as negative (healthy).

V. CONCLUSION AND FUTURE WORK

A pandemic is happening more and more day by day in recent times. Various sectors such as agriculture, business, finance are rapidly deteriorating at this time. Many people lose their jobs and are unable to manage new things due to the trapped situation. This condition occurs because the disease can be transmitted from person to person in a very short time. Undetectable cases of COVID-19 can spread the disease to their community quickly. Therefore, early detection is the only solution to detect COVID-19 cases and take appropriate action as far as possible. To perform this function, a relatively inexpensive tool is needed that can detect COVID-19 cases as quickly as possible in a large number of cases. Our proposed CNN model can detect the positive and negative effects of COVID-19 and make people trustworthy. This model can help people get COVID charges early. This study will now detect COVID-19 by investigating X-ray images of patients' chests and reducing the spread of the disease early. We proposed both machine reading and in-depth study based on investigating the common open source of pneumonia, viral pneumonia and COVID-19 patient images to predict COVID-19 automatically in the short term. Our results show that CNN scored high with 94.01% accuracy and 97.01% accuracy, while many of the deep learning methods represent good results as standard dividers. Our approach may be helpful in clinical practice to detect COVID-19 cases early and to prevent future public outbreaks.

1. REFERENCES
2. Asmaa Abbas, Mohammed M Abdelsamea, and Mohamed Medhat Gaber. “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network”. In: arXiv preprint arXiv:2003.13815 (2020).
3. Md Zahangir Alom et al. “COVID MTNet: COVID-19 Detection with Multi-Task Deep Learning Approaches”. In: arXiv preprint arXiv:2004.03747 (2020).
4. Ioannis D Apostolopoulos and Tzani A Mpesiana. “Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks”. In: Physical and Engineering Sciences in Medicine (2020).
5. World Health Organization, "Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020," World Health Organization, World Health Organization2020.
6. S. U. K. Bukhari, S. S. K. Bukhari, A. Syed, and S. S. H. SHAH, "The diagnostic evaluation of Convolutional Neural Network (CNN) for the assessment of chest X- ray of patients infected with COVID-19," medRxiv, 2020.
7. H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, et al., "Radiological findings from 81 patients withCOVID-

19 pneumonia in Wuhan, China: a descriptive study," The Lancet Infectious Diseases, 2020.

1. Linda Wang and Alexander Wong. “COVID- Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images”. In: arXiv (2020), arXiv–2003.