COVID-19 Data Analysis using Chest X-ray

Ishtiaque Ahmed, Manan Darda, Neha Tikyani, Rachit Agrawal, Manjusha Joshi

Abstract: The COVID-19 pandemic has caused large-scale outbreaks in more than 150 countries worldwide, causing massive damage to the livelihood of many people. The capacity to identify contaminated patients early and get unique treatment is quite possibly the primary stride in the battle against COVID-19. One of the quickest ways to diagnose patients is to use radiography and radiology images to detect the disease. Early studies have shown that chest X-rays of patients infected with COVID-19 have unique abnormalities. To identify COVID-19 patients from chest X-ray images, we used various deep learning models based on previous studies. We first compiled a data set of 2,815 chest radiographs from public sources. The model produces reliable and stable results with an accuracy of 91.6%, a Positive Predictive Value of 80%, a Negative Predictive Value of 100%, specificity of 87.50%, and Sensitivity of 100%. It is observed that the CNN-based architecture can diagnose COVID-19 disease. The parameters’ outcomes can be further improved by increasing the dataset size and by developing the CNN-based architecture for training the model.

Keywords: COVID-19, CNN, Chest X-ray, ResNet18, Radiograph Images, Deep Learning

I. INTRODUCTION

The new coronavirus (SARS-CoV2) has spread to Wuhan, China, and other countries since December 2019. By April 18, more than 2 million infected people and more than 150,000 deaths had been recorded worldwide [1]. Early diagnosis is essential to quickly isolate suspected individuals due to the new COVID-19 or lack of vaccines and lessen the danger of infection in the general population.

Fever, cough, myalgia, or fatigue are common initial symptoms in patients [2]. Acute respiratory syndromes, which can lead to pneumonia and death, are more severe symptoms that patients often experience [3],[4].

As a result, coronavirus is also known as acute pneumonia. Pneumonia and a chest X-ray are frequently associated with diagnoses made in patients. Due to the rapid development of advanced technology, many studies are being conducted on the coronavirus (COVID-19) to use artificial intelligence (also known as machine learning) for the early detection of patients with this disease [5].

The COVID-19 pandemic is causing extreme harm to the respiratory system and other human organs. Therefore, the medical imaging features of chest X-rays have helped quickly detect COVID-19. Medical imaging modalities, such as X-rays can be used to obtain imaging features of the chest. Radiography has many benefits, including:

1. Forming a 3D view of the organ.
2. Quick checking the disease.
3. The location of the disease.

Additionally, the equipment required to obtain these CT images is already insufficient supply in all countries. Therefore, chest X-rays have drawn the attention of researchers to testing for COVID-19. To use chest radiographs to detect COVID-19 accurately and quickly, radiologists’ guidance is required [6]. Treating COVID-19 speedily and accurately is a difficult task for healthcare providers. However, the shortage of traditional COVID-19 testing kits is a significant problem. To reduce human involvement in chest imaging to detect diseases, a self-diagnostic model is needed to see COVID-19 using imaging methods.

A. Using Chest X-rays to detect COVID-19

Since the recent surge in COVID-19 infections worldwide, many alternative testing methods have been established to detect suspected COVID-19 cases. There are, however, a limited number of open-source applications that use chest X-ray images [7]. There is a limited amount of data on exposure to COVID-19 chest X-rays. ResNet18 is the only way to provide an open-source actively maintained tool that can detect not only COVID-19 but also other cases of pneumonia with high detection sensitivity for COVID-19. COVID-Net learns architectural design using machine-driven design surveys, starting with initial design prototypes and requirements [8].

B. What are we extracting from Chest X-Rays?

Chest radiography imaging, used as a standard tool for pneumonia diagnosis, is simple to use and provides quick results.

• X-ray has a high sensitivity for diagnosing COVID-19.
• The Multilobar involvement and the opacity of the peripheral airspace can be seen in the chest image. Mixed attenuation with ground glass (57%) is the most commonly reported opacity (29%) [9].
• In the early stages, a ground glass pattern can be seen around the pulmonary vessels’ edges, which can be difficult to see visually [10].
COVID-19 has also been linked to asymmetric patchy or diffuse airspace opacities.

- Machine Learning and artificial intelligence (AI) solutions can be constructive in solving these problems.

C. Using Residual ConvNet – ResNet18 to detect COVID-19

The pre-prepared ResNet18 model, which was prepared on the ImageNet dataset, utilised in this investigation was one of the models [12]. ResNet is perhaps the most famous CNN structure, and it was the victor of the 2015 ImageNet contest. It permits smoother gradient flow for a more proficient preparation. ResNet's central idea is to present an "identity alternate way" association that sidesteps at least one layer [13]. This makes it a lot simpler to refresh gradients in these layers, permitting the network to give direct ways to early layers of the network. The accompanying figure shows the overall outline of the ResNet18 model and how it is utilised to determine COVID-19 [14].

II. METHODOLOGY

We analyse performance for COVID-19 detection using ResNet18 in a convolution network with a COVID X-ray-3k dataset.

- First, we would be transforming the images Resolution of the X-Ray as per ResNet18 format.
- As the images in the dataset are not used directly, and we would be providing augmented Images to the model.
- To solve this, the CCN model based on transfer learning with pre-trained data augmentation is run.
- This can be done by training the dataset with the model.
- The model predicts a probability score for each image. This shows the possibility that the image can be detected as COVID-19.
- By comparing this probability to the cut-off threshold, we can obtain a binary label indicating whether the image is COVID-19.

In the planned work, the vast requirement for an outsized variety of COVID-positive Chest Radiograph datasets is settled utilising a technique known as stationary wavelet-based data augmentation [15]. The prepared model set includes the aggregated images of the multi-scale discriminant highlights prepared and merged to identify the binary class names COVID-19 and non-COVID-19 [16]. The figure below shows the different levels of description of the three-stage COVID-19 detection method by lung chest radiographs:

![Block Diagram](image-url)
A. Phase-1 (Image Augmentation)

In this study, a chest radiograph image should be viable with the recently prepared transfer learning-based model. Therefore, the pre-processing stage is: Changing the type of input image data, adjusting the input image size, and normalising the input image. [17],[18]. However, the number of chest X-rays available for positive COVID-19 cases is limited. To solve this problem, a pre-trained CNN model based on augmented transfer learning was implemented.

The proposed method was performed based on chest X-ray images of the lungs of normal and positive COVID-19 patients. [19]. The database contains 349 positive chest X-rays, positive COVID-19 chest X-rays from 216 patients, and 397 chest X-rays from non-COVID patients. The input chest radiograph image has various image sizes and formats such as JPEG, PNG.

B. Phase-2 (Transfer Learning Models)

The COVID-19 detection model depends on a previously trained conversion learning method, which divides chest X-rays into two categories: a) COVID-19 and b) non-COVID-19. ResNet18 model is used for binary classification, which improves training and adjusts image size based on the similarity with pre-trained multiple CNN models. 224x224x3 is the required dimension for the transfer learning model ResNet18.

The pre-training model can classify chest radiographs according to the class labels (COVID-19 and non-COVID) assigned to the training dataset. [20]. By retraining our pre-trained model by updating a fully connected layer on the input scaling dataset to classify radiograph images. Based on transmission learning, the following training parameters for the CNN model were selected: a) the 'sgdm' (Stochastic Gradient Descent with momentum) optimiser used, b) a small batch size of 64, b) training was conducted from up to 50 epochs, c) value of validation frequency is set to 3 and d) 3e-4 should be the pre-set predefined learning rate for the training. It is also used in the deeper layers of the best performance models to detect abnormalities in chest radiographs of positive COVID-19 cases [21].

C. Phase-3 (Using Deep Layer to Locate Abnormalities)

In this proposed method, examining the already trained network layer and the activation of various channels are done. Each layer of the pre-trained CNN network contains a channel (consisting of many 2D arrays). Deeper layers and convolutional first layer (conv1) output activations are used to identify abnormalities in positive COVID-19 chest X-ray images through optimised training of the transfer learning model. The functionality acquired by the network was assessed based on COVID-19 chest X-ray images with an input activation zone. Activation is obtained in a 3D matrix with three dimensions that indexes the channels of the layer to be inspected. Ideally, conv1 will detect colour and edge-like features, and deeper layers will detect more complex features.

Regardless of many imaging modalities, we believe that chest X-rays have a sensitivity that is not highly optimised for significant clinical discoveries. Radiograph images are a common way for doctors to diagnose pneumonia. The X-ray imaging system is a significant part of the world's medical services. The ease of use of the X-ray machine takes time to detect COVID-19 cases without a screening workbench or kit. There may also be specific cases in which the patient takes an image for other reasons and discovers a potential implication of the scan result COVID-19 [22]. Radiographs pictures are pictures of the patient's chest taken within one year prior and then afterwards the COVID-19 pandemic. The consequences of the radiograph images firmly propose that the impact can be seen from the lungs, especially in the peripheral and subpleural distribution and in the lower lobes and posterior segment, even in the beginning stages of COVID-19. Over time, the spread of this disease becomes more increasingly widespread. In any case, the most significant challenge experienced here is that it requires some time and medical specialists in the field to look at every radiograph image and extract significant discoveries.

As a result, doctors need computer assistance to help detect cases of COVID-19 using X-ray images. In today's situation, where hundreds of thousands of people must be checked every day for the deadly COVID-19 virus, they must use automated, reliable, and accurate computer-assisted methods to determine the presence of the infection. Deep learning techniques in computer-aided technology have contributed to medical image analysis in the most advanced way and have shown excellent performance. Therefore, the current study proposes a dip-running-based two-step method for detecting and classifying cases of pneumonia using radiograph images [23].

Fig.3. Flow Chart

III. RELATED WORKS

Broad exploration work is continuing for arranging Covid-19 patient picture information. Many scientists have proposed a unique Deep Learning Model for organising Chest X-ray images while reviewing CT images [24].

The absence of information in clinical imaging drove us to investigate methods of extending picture datasets. In current navigation, we focus on improving Covid-19 awareness. An Auxiliary Classifier Generative Adversarial Network (ACGAN) based model is built by blending all standard CXR pictures.
Radiology and chest X-rays in a questionnaire survey placed in the journal have surpassed research centre testing to detect new Covid infections in 2019 [25]. Radiographic abnormal recurrences escalated immediately after the onset of adverse events and peaked at the time of the disease.

The scientists concluded that Chest X-rays should be used as basic detection technology for COVID-19. Specifically, chest radiograph images have different advantages, such as being fast and easy to access, easy to use and helping to determine the priority of suspected coronavirus patients quickly. Due to the lateness of the pandemic, only a predetermined number of chest x-ray images are available for research. Therefore, Covid GAN is created to generate fake preparation information for CNNs. The age of fake information is compelling because of tiny datasets and when the information incorporates touchy data. GANs can blend pictures without any preparation from a particular classification and deliver agreeable outcomes when joined with different techniques.

Many exploration ventures and advancements identified with COVID-19 have been proposed. This paper, be that as it may, as far as we could know, is the first to introduce a GAN engineering for development in COVID-19 identification.

A. AI-Based X-ray Images Analysis

The use of AI for ahead of time discovery, analysis, observation, and creation of immunisations for COVID-19 was intricately examined in a few examinations in the writing that misused different profound learning procedures on X-ray information show sensible execution [26]. A model for early recognition of COVID-19 called DarkCovidNet has been proposed. This model used a layer of 17 convolutions to run double-class and multi-class batches, typically including cases of COVID and pneumonia. Although the model detailed a typical accuracy of 98.08% for pair ordering and 87.02% for multi-class batches, DarkCovidNet's reproduction using different data sets resulted in overtraining and much lower accuracy when unilateral test information was introduced into the model. Appear. Several papers applied a heavy learning model to CT filter images to discriminatingly screens X-ray information COVID-19 highlights. Ardakani et al. Differentiate between cases of with and without COVID-19 in the implemented cutting-edge CNN models utilised, e.g., ResNet18, ResNet-50, ResNet-101, SqueezeNet, VGG-16, VGG-19, MobileNet-V2, GoogleNet. According to their analysis, deep learning may be seen as a practical step in distinguishing between COVID-19 and X-ray [27]. To avoid under-guessing or overfitting due to the lack of an inaccessible data set for the COVID-19 example, we used a GAN model to generate manufacturing information and achieved a die coefficient of 0.837. The relevance of GANs for the combination of COVID-19 radiographic information can be found in broader GAN applications for a variety of clinical information. The study distinguished different exceptional properties of GAN like area transformation, information growth, also, picture-to-picture interpretation that urged specialists to receive it for picture reproduction, division, identification, characterisation, and cross-methodology blend for different clinical applications.

B. Role of Imaging Modalities

Most nations expected to take measures to react to the sudden and fast eruption of COVID-19 within a respectfully short period. Radiology divisions have begun to zero in on readiness instead of demonstrative capacity after adequate information was accumulated for COVID-19 [28]. The trial showed similarities to COVID-19 with various infections brought about by other Covid strains, such as severe acute respiratory syndrome (SARS). The significance of using CT scans to track the lung status of recovering Covid patients was similarly referenced in investigations [29]. The chest imaging strategy was an essential strategy for distinguishing COVID-19 by identifying bilateral nodules and fringe frosted glass opacities on radiographs of the lungs.

IV. RESULT

The cautious assessment of clinical models is a significant undertaking, as numerous significant choices might be impacted by the yields given by such models. An appropriate assessment thinks about exactness as well as thinks about different variables as portrayed beneath

1) Confusion Matrix

The confusion matrix is a summary of the prediction results of classification problems. Summarises the number of correct and incorrect predictions.

It is handy for measuring actual values, predicted values, sensitivity, specificity, and accuracy. Table 1 shows the expected outcomes of the model using the confusion matrix table.

Table 1. Confusion Matrix Table

<table>
<thead>
<tr>
<th>Predicted Normal/Pneumonia</th>
<th>Normal/Pneumonia</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Normal/Pneumonia</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>Predicted Covid</td>
<td>0</td>
<td>12</td>
</tr>
</tbody>
</table>

2) Positive Predictive Value (PPV)

It predicts that the patient will experience an infection because the model expected the information test to be accurate, as shown in condition 1.

\[ PPV = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \]  

3) Negative Predictive Value (NPV)

It determines the likelihood that a patient doesn't experience the ill effects of an illness given that the model has anticipated the information test as unfavourable, as demonstrated by condition 2.

\[ NPV = \frac{\text{True Negative}}{\text{True Negative} + \text{False Negative}} \]  

4) Sensitivity

It is the range of actual specific cases expected to be positive in all information tests, as shown in condition 3.
image method needs to be involved in different methods (e.g., internal heat level, ECG, MCG, diabetes), the CNN model needs to be customised. Level renal capacity, etc.) and patient boundaries (age, gender, nationality, travel history, presence or absence of contact, etc.) to identify the nature of COVID-19 internally and externally. The definition of AI models related to the depiction of multi-module information is needed later to decipher and clarify the characteristic evaluation results.

VI. CONCLUSIONS AND FUTURE SCOPES
The coronavirus is affecting the health of the world's population at a shocking rate. Early intervention in patients and prevention of infection risk is necessary for early treatment. Given the impact of the coronavirus on inhaled tissues, chest X-rays are a critical part of the assessment of lung abnormalities and patterns for disease to aid in the detection and determination of severity. Also, against this pandemic, the use of portable devices to hold chest x-rays allows for limiting disease management issues, first addressing basic radiographic methods for more viable segmentation than contaminated patients. Therefore, there is a risk of cross-contamination.

Therefore, chest radiographs obtained from patients with and without Covid-19 were used. These images are sort using the ResNet18 transfer learning model. For the model's training, a set of 3000 chest X-ray images was classified into three classes: "COVID-19", 'normal' and 'viral pneumonia'. The change of grouping parameters using the CNN technique indicates that it is relatively robust. A subset of the preparation information also provides the appropriate characteristics by leveraging the tremendous ideal preparation information. From the results we can say that to correctly diagnose the COVID-19 disease the CNN-based architecture has the potential. In improving detection accuracy transfer learning plays an important role.

In the future, more effective deep learning models can be implemented. In addition, it can work with larger information data sets. Our future bearings will be centre around applying this strategy continuously in clinical information data sets. Our future bearings will be centre around applying this strategy continuously in clinical information and improving the exactness of our research study. Future work may include the development of a new CNN-based architecture for detecting COVID-19 and other diseases in the medical field.

REFERENCES


AUTHORS PROFILE

Ishiaque Ahmed, Student, MBA-TECH, NMIMS MPSTME, Mumbai, Maharashtra, 400056
Area of Specialization- Technical Stream: Electronics and Telecommunication, Management Stream: Finance, Past Experiences: Raymond Ltd. (Summer Intern), Being Digital (Business Research Intern), Publications: Blockchain: A New Safeguard to Cybersecurity (Springer), A Case Study on Rapidio

Manan Darda, Student, MBA-TECH, NMIMS MPSTME, Mumbai, Maharashtra, 400056

Neha Tikyan, Student, MBA-TECH, NMIMS MPSTME, Mumbai, Maharashtra, 400056
Area of Specialization- Technical Stream: Electronics and Telecommunication, Management Stream: Marketing, Marketing Enthusiast, Past Experiences: Technology Twist (Web Development Intern), Glossaread Technologies Pvt. Ltd. (Content Coordination), NetCom Learning (Data Research Intern)

Rachit Agraval, Student, MBA-TECH, NMIMS MPSTME, Mumbai, Maharashtra, 400056
Area of Specialization- Technical Stream: Electronics and Telecommunication, Management Stream: Marketing Past Experiences: Medoz Pharmaceuticals (Intern), BHEL Bhopal (Intern)

Dr. Manjusha Joshi, Assistant Professor, Electronics and Telecommunication Engineering Department, NMIMS MPSTME, Mumbai, Maharashtra, 400056
Specialization: Biomedical signal processing, Neural network B.E., M.E., PhD

Teaching Experience: 20 Years, Research and Publication: Publications – Conferences (International – 05) and Publications – Journals (International-11)

Honours and Awards: Best paper at international conference – ‘artificial intelligence in healthcare’ in 2015, Invited as speaker at ‘Emerging Trends in Pharma Sciences’ in 2018, faculty training program at Pharmacy College (SPTM) Shirpur, Reviewed 20 research paper for biomedical signal processing and control, Elsevier