

# Predicting COVID-19 Pneumonia Severity based on Chest X-ray with Deep Learning

Swati Shekapure  
Guide, Marathwada Mitra  
Mandal's College of  
Engineering, Pune

Nikita Pagar  
Student, Marathwada Mitra  
Mandal's College of  
Engineering, Pune

Bhagyashree Kulkarni  
Student, Marathwada Mitra  
Mandal's College of  
Engineering, Pune

Dinesh Choudhary  
Student, Marathwada Mitra Mandal's College  
of Engineering, Pune

Priti Parkhad  
Student, Marathwada Mitra Mandal's College  
of Engineering, Pune

## ABSTRACT

Pneumonia is an infectious disease that affects one or both lungs in the human body commonly caused by bacteria called *Streptococcus pneumoniae*. It is an infection of microscopic particles in the air sacs of the lungs, called alveoli. Chest X-Rays are used to diagnose pneumonia and which needs an expert radiologist for evaluation. This may vary over time from practitioner to practitioner. This is based upon the person's experience too. Therefore, an automated system is required that can help patients to diagnose pneumonia without any of these constraints. We propose an image-based automated system that detects pneumonia diseases using Artificial intelligence. The system will be making the use of computational techniques for analyzing, processing, and classifying the image data predicated upon various features of the images. Unwanted noise is filtered and the resulting image is processed for enhancing the image. Complex techniques are used for feature extraction like the Convolutional Neural Network (CNN) followed by classifying images based upon various algorithms. The diagnosis report is generated as an output that also contains a severity score. This system will generate more precise results and will provide them faster than the traditional method, making this application more efficient and dependable. This application can also be used as a real-time teaching tool for medical students in the radiology domain.

## Keywords

Image Processing, Artificial Intelligence(AI), Neural Network, Deep Learning, COVID-19, Viral pneumonia.

## 1. INTRODUCTION

Coronavirus is caused by a severe acute respiratory syndrome coronavirus 2 i.e SARS-CoV-2, which was first found in Wuhan City, Hubei Province, China. The Chinese authorities reported an increasing number of cases of pneumonia, which they later identified the causative organism to be a coronavirus. Various deep learning algorithms have obtained much attention for classification of the disease to analyze medical images. Also various features are learned and extracted from the pre-trained CNN model on large-scale datasets are useful in image classification tasks. In this process, we need to evaluate the functionality of pre-trained CNN models utilized for feature-extraction followed by different classifiers for the classification of abnormal and normal chest X-Rays. We analytically determine the CNN

model for this purpose. Statistics show that the results obtained to demonstrate that pre-trained CNN models worked along with supervised classifier algorithms can be very beneficial in analyzing images used to predict Pneumonia. This system can also redirect the images to an expert in case the accuracy of the classification for a particular image is found to be very low. This solution is targeted for use by the general public for remote self-diagnosis. Chest X-rays provide a non-invasive tool to monitor the progression of the disease. Images from COVID-19 pneumonia database backtracked by three blinded experts in terms of the extent of lung involvement as well as the degree of opacity.

## 2. LITERATURE SURVEY

The use of artificial intelligence-based systems is very common in detecting those caught in the COVID-19 epidemic. As given in Table, there are many studies on this subject in the literature. In binary classification, it is common to distinguish COVID-19 positive from COVID-19 negative. The researchers used 70% data for the training, 10% for validation, and 20% for the test. As a result, they obtained 94.40% accuracy over test data with the CNN model they suggested and proposed a two-class study using limited data. They reported their performances by dividing the dataset at different training and testing rates. They achieved the highest accuracy of  $94.65 \pm 2.1$  at 70% training - 30% testing rates. In their study, they set the CNN hyper-parameters using multi-objective adaptive differential evolution. Afshar et al. conducted their studies using a method called COVIDCAPS with multi-class (Normal, bacterial pneumonia, viral pneumonia, and COVID-19) studies. They achieved 95.7% accuracy with the approach without pre-training and 98.3% accuracy with pre-trained COVID-CAPS. However, although their sensitivity values are lower than general accuracy, they detected the without pre-training and 98.3% accuracy with pre-trained COVID-CAPS as 90% and 80%, respectively. From their point of view, considering that this pandemic period affects the whole world, there is a serious increase in the work density of radiologists. In these manual diagnoses and determinations, the expert's tiredness may increase the error rate. It is clear that decision support systems will be needed in order to eliminate this problem. Thus, a more effective diagnosis can be made. The most important issue that restricts this study is to work with limited data. Increasing the data, testing it with the data in many different centers will enable the creation of more stable systems. In future studies,

the features will be extracted using image processing methods on Xray and CT scan images. From these extracted features, the features that provide the best separation between classes will be determined and performance values will be measured with different classification algorithms. In addition, the results will be compared with deep learning models. Apart from this, the results obtained from this study will be tested with data from many different centers. In a future study, studies will be conducted to determine the demographic characteristics of patients and COVID-19 possibilities with artificial intelligence-based systems. In another study, Zhang and his colleagues performed binary and multi-class classifications containing 106 COVID-19 data. They found a detection accuracy of 95.18% with the CAAD model. Apostolopoulos et al. obtained an accuracy of 93.48% using a total of 224 COVID-19 data with the VGG-19 CNN model for their 3-classes (COVID19 -bacteria-normal) study. Narin et al. used 50 COVID-19 / 50 Normal data in their study, where they achieved 98% accuracy with ResNet50. In many studies in the literature, researchers have studied a limited number of COVID-19 data.

### 3. MATERIALS AND METHODS

'covid19-pneumonia-normal-chest-x-ray-pa-dat aset' is a freely available dataset of around 5000 images gathered and labeled by Amanullah Asraf.

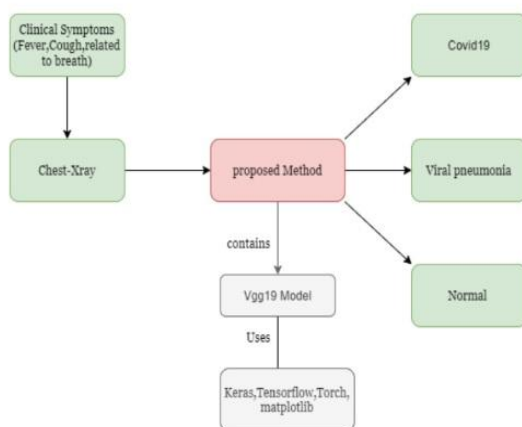


Figure 1: System Overview

### 4. SYSTEM ARCHITECTURE

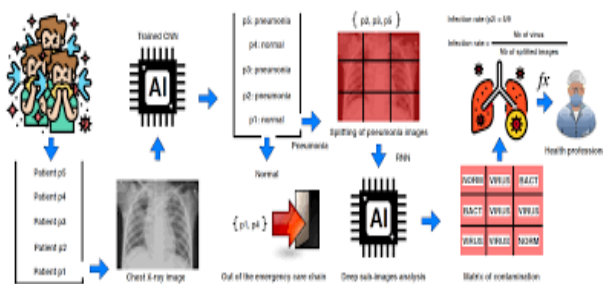


Figure 2: System Architecture

The overall system architecture shows the different modules in the system and how they are connected.

#### 4.1. Image Acquisition

In the image acquisition, the images for the system are

acquired by appropriate scanning of X-rays in the dataset. The input image is obtained by camera or through some scanner. The input image may be in gray color.

#### 4.2. Preprocessing

The preprocessing consists of a series of operations performed on the scan input image, which includes background noise reduction, image restoration, filtering, etc. Various algorithms may be used for this purpose.

#### 4.3. Segmentation

The segmentation algorithms partition an image into sets of pixels or regions. The need for partitioning the data is to better understand what the image represents.

#### 4.4. Classification using CNN

This module will classify the image to actually diagnose the covid pneumonia. It will use a CNN to perform the classification.

#### 4.5. Recommendation System

This system will recommend the nearby doctors or specialists and any available home cures or remedies or preventive measures.

#### 4.6. Report Generation

Based on the results from the classification model, a report will be generated which will include the disease detected, nearby radiologist, common cures, and prevention models.

### 5. TOOLS AND TECHNOLOGIES

Tools and Technologies that will be used in the application are as follows:

#### 5.1. Image Processing

Python & Keras

#### 5.2. Image Classification

Python, Tensorflow

#### 5.3. Web Application

HTML, CSS, JavaScript, Flask, Bootstrap

#### 5.4. Other Required Technologies

Git, Github, AWS

### 6. USER INTERFACE

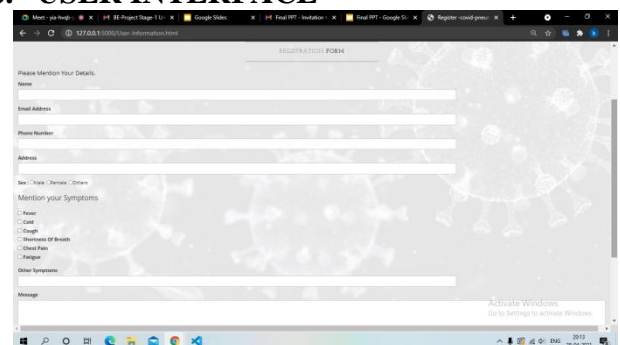


Figure 3: Registration Form

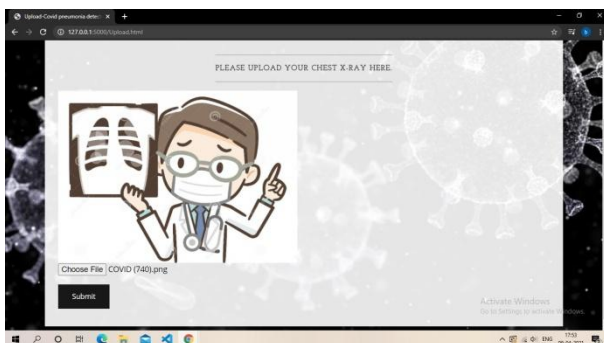


Figure 4: Uploading chest x-ray image

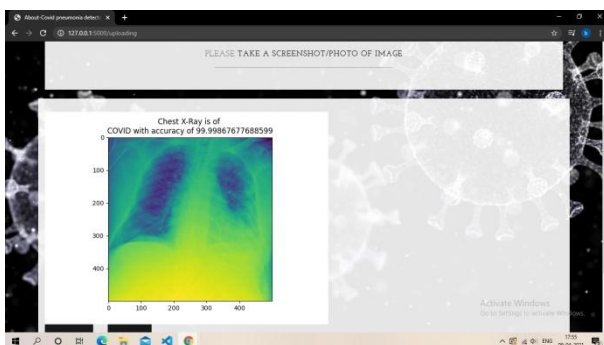


Figure 5: Output

## 7. ADVANTAGES

1. CNN is used which extracts features accurately.
2. Non-invasive
3. Diagnosis of the disease is provided without any delay in remote areas.
4. Ease of accessibility
5. Provides immediate relief measures and precautions
6. Solves the issue regarding the small number of radiologist catering to a larger population
7. Time and money will be saved.
8. The user can access the system remotely from the comfort of his/her home.

## 8. LIMITATIONS

1. Requires an internet connection.

2. The currently developed AI-based model is a little bit slow in the classification of chest X-rays.
3. Image quality might lose during the scanning process.
4. The final report will be generated based on the x-ray image so it should be clear otherwise the report will not generate properly.

## 9. CONCLUSION

AI-based classification can help rapid diagnosis of COVID-19 pneumonia diseases. The models developed by us are proof of the concept that cost-effective, user-friendly, and non-invasive AI-based methods can be developed for COVID-19 pneumonia. The AI-based models developed by us may be evaluated for its use in clinics, as diagnostic or clinical management of patients. Also, in the future, with the availability of more and more images, representing diverse cases, the efficiency of the models may be scaled up.

## 10. REFERENCES

- [1] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri and Ronald M Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases", Computer Vision and Pattern Recognition (CVPR) 2017 IEEE Conference on, pp.34623471, 2017.
- [2] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine, vol. 121, p. 103792, 2020.
- [3] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: concepts, methods, challenges, and future directions," IEEE Access, vol. 7, pp. 99540–99572, 2019.
- [4] Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology 2020 Feb 26;200642 [FREE Full text] [CrossRef] [Medline].
- [5] X. Xu, X. Jiang, C. Ma et al., "A deep learning system to screen novel coronavirus disease 2019 pneumonia," Engineering, 2020.