COVID 19 Detection Using Chest X-ray Images applying Deep Learning

Created by: Manasij Haldar

Main research activities

Developing models and algorithms for the analysis of biomedical images:

- To support the medical diagnosis with Deep Learning;
- To allow large-scale image analyses.

Objective

- During a pandemic, such as COVID-19, a timely and precise diagnosis is critical. It improves patient outcomes and relieves burden on health-care systems that are dealing with a growing rate of patients.

Covid 19

COVID-19 is a virus that affects the lungs and causes tissue damage in affected people. Some people may not experience any symptoms in the early stages, whereas the majority of patients experience fever and cough as their primary symptoms. Other probable secondary symptoms include body aches, a sore throat, and a headache.

COVID-19 disease is currently on the rise due to a lack of early screening methods. In 2020, a large number of individuals perished from this sickness all across the world.

COVID 19 SCREENING

PCR

The polymerase chain reaction (PCR) is the current preferred approach for diagnosing COVID-19 (PCR). However, some of the worst-affected locations are unable to obtain enough kits to fulfill demand, and many countries are unable to conduct tests due to a lack of lab facilities.

Deep Learning Approach

Deep learning algorithms could be employed in this case to identify infected individuals using chest X-ray scans, which are readily available around the world. This approach could be employed in situations where PCR diagnostics are currently unavailable.

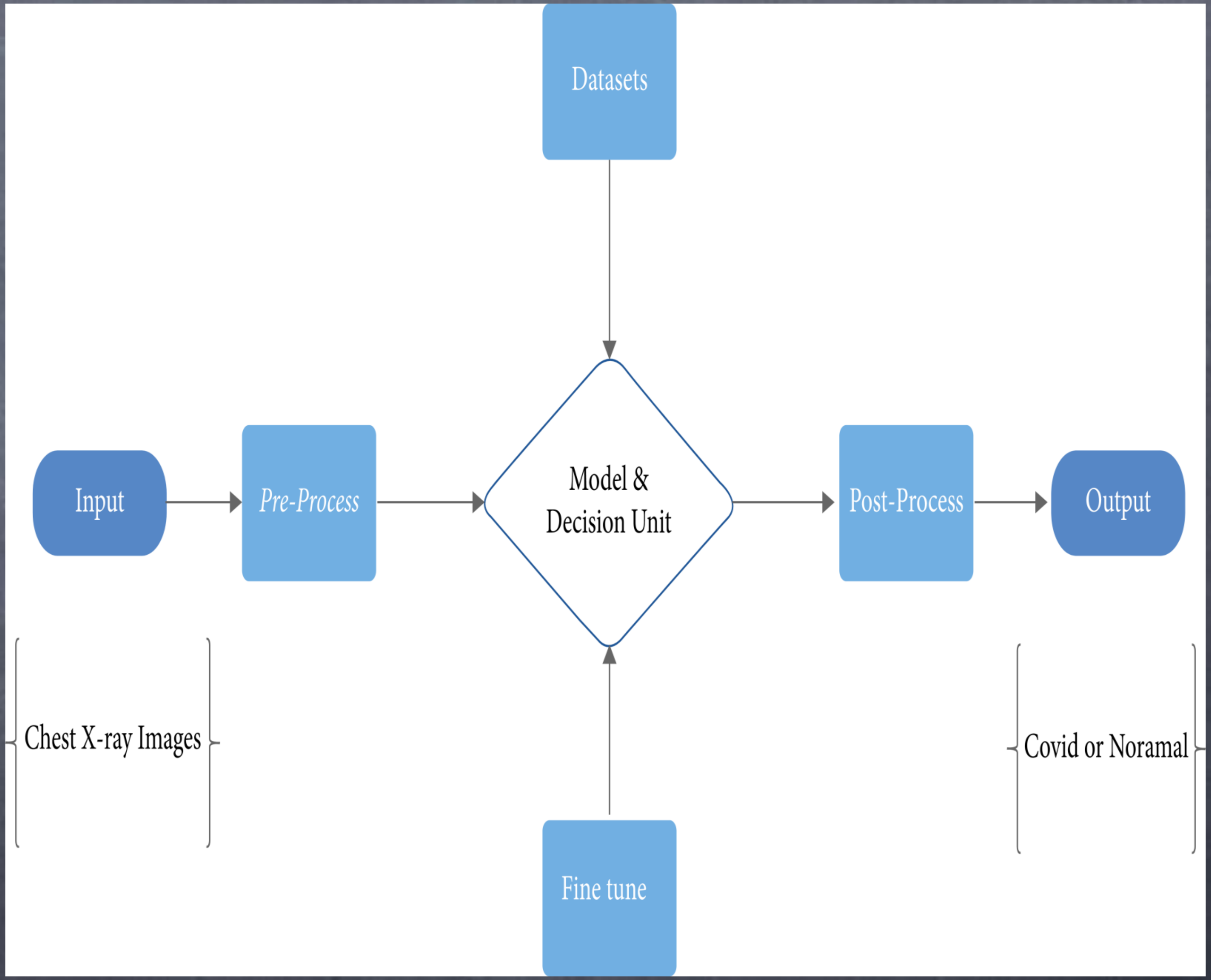
Deep learning for X-ray analysis might drastically cut the time it takes to diagnose patients, with an AI model processing up to 200 images in the time it takes a radiologist to analyze one.

Dataset

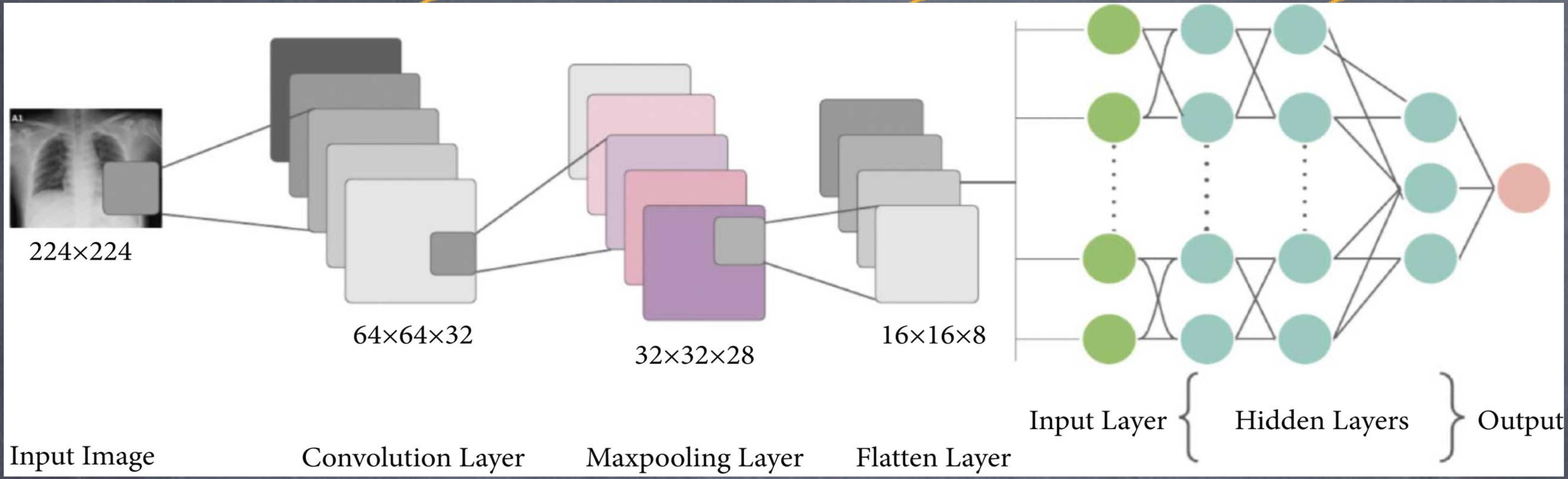
We propose using a deep convolutional neural network trained on COVID-19 and pneumonia images, as well as a fresh dataset comprising COVID-19 and pneumonia images. Both are open to the public via GitHub and Kaggle , respectively. COVID-19 cases are represented by the chest X-ray or CT images available on GitHub. It was made by putting together medical photos from publicly accessible sources and publications. There are 204 COVID-19 X-ray pictures in this dataset. The Kaggle dataset, on the other hand, was produced for a pneumonia detection challenge. Bounding boxes surround sick lung regions in the photos. Without the boundary boxes, the samples are negative and show no signs of pneumonia. The presence of bounding boxes in samples indicates the presence of pneumonia.

Block Diagram

The input is given as a CXR image of a dataset, which has two subsections: COVID-19 patients and normal patients. Before fitting the model, this system underwent preprocessing, such as loading images of a particular size, splitting dataset, and data augmentation techniques.



System Architecture

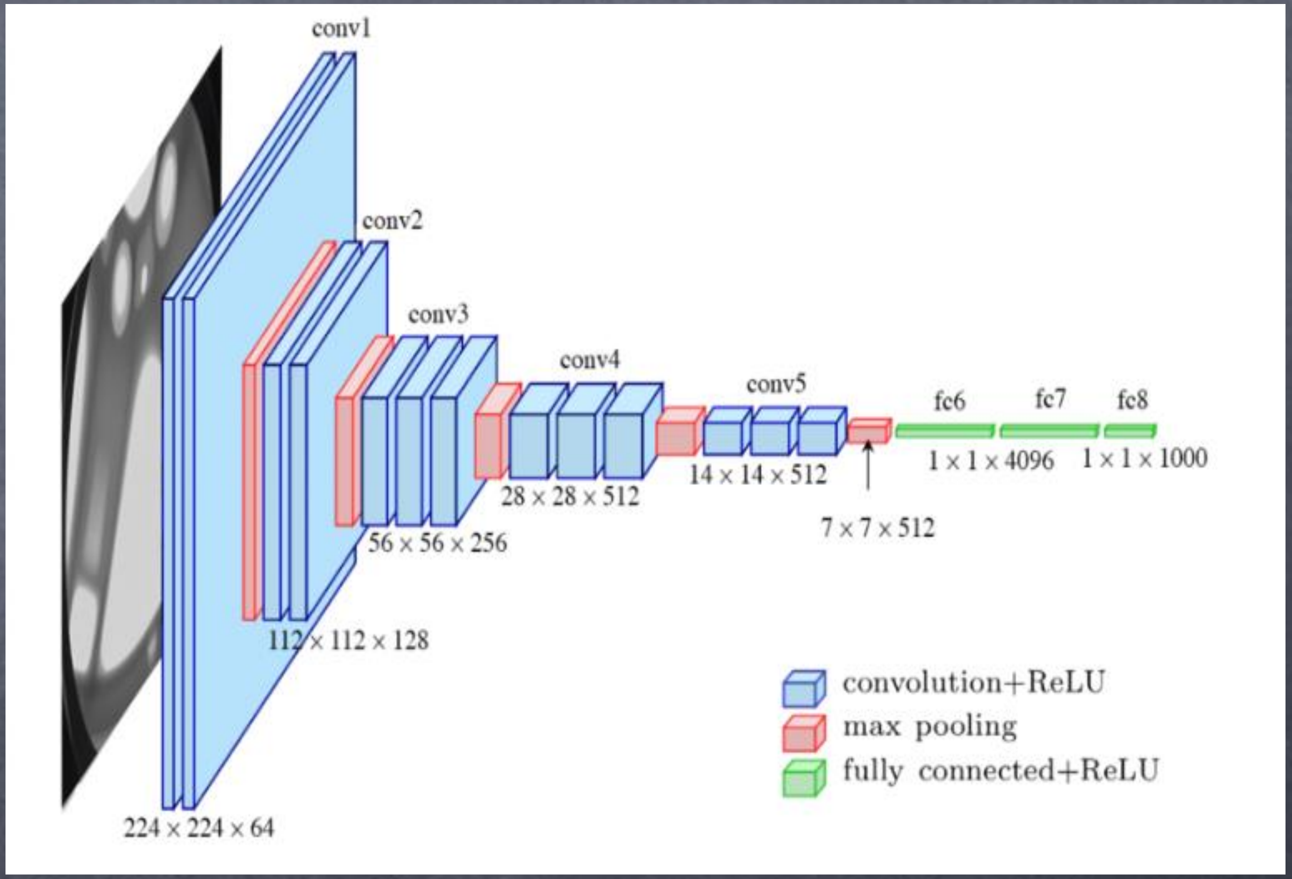


In this architecture, the input is a CXR image, and the output is a prediction of the image. In this case, it will predict whether the image is COVID-19 affected. The input shape is 224 × 224, and there are three channels. In the first two layers of the designed architecture, the filter size is 32 with padding, kernel size of 3, and activation function as ReLU. Thereafter, there is the first maxpooling layer, which has a pool size of 2 and strides of 2. The following layer is a flat layer that converts pooled features into a single column. Finally, two dense layers were formed. The first one has ReLU as an activation function, and the least dense layer's activation function is softmax.

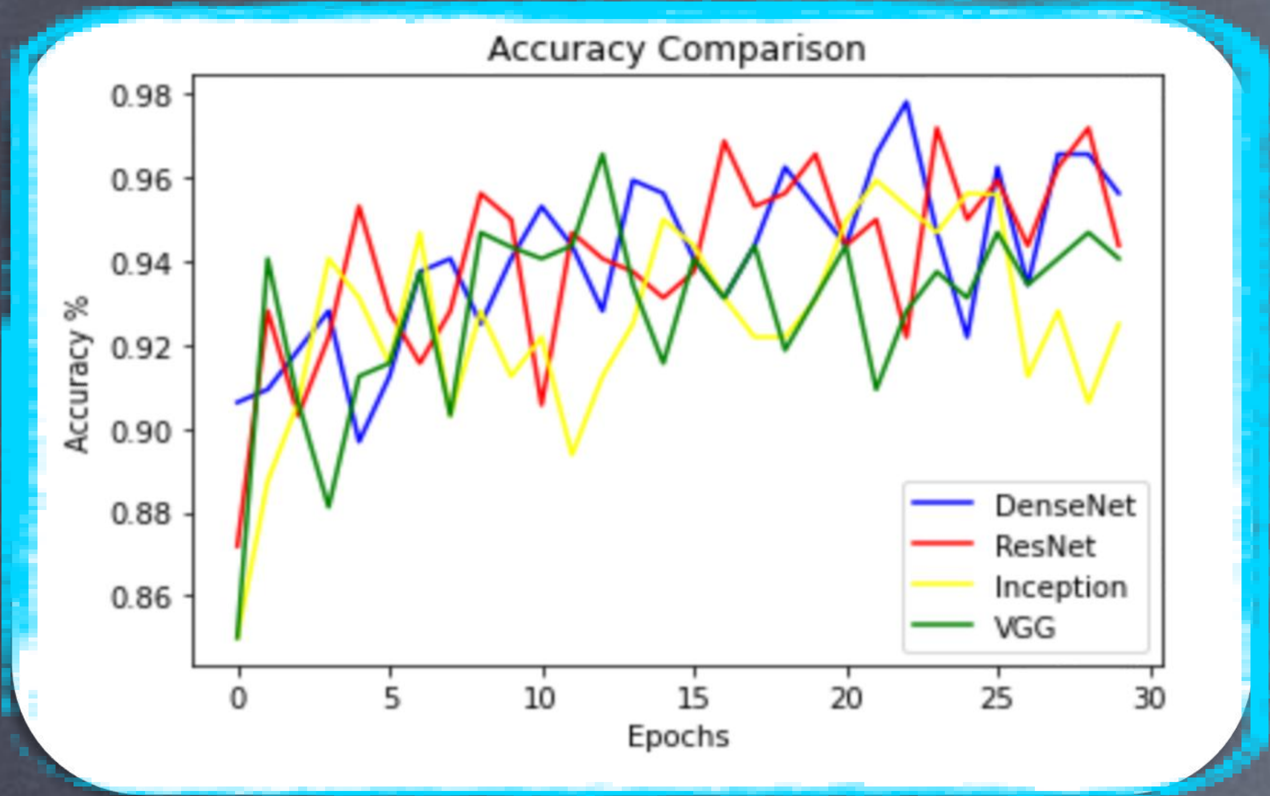
Vgg-16

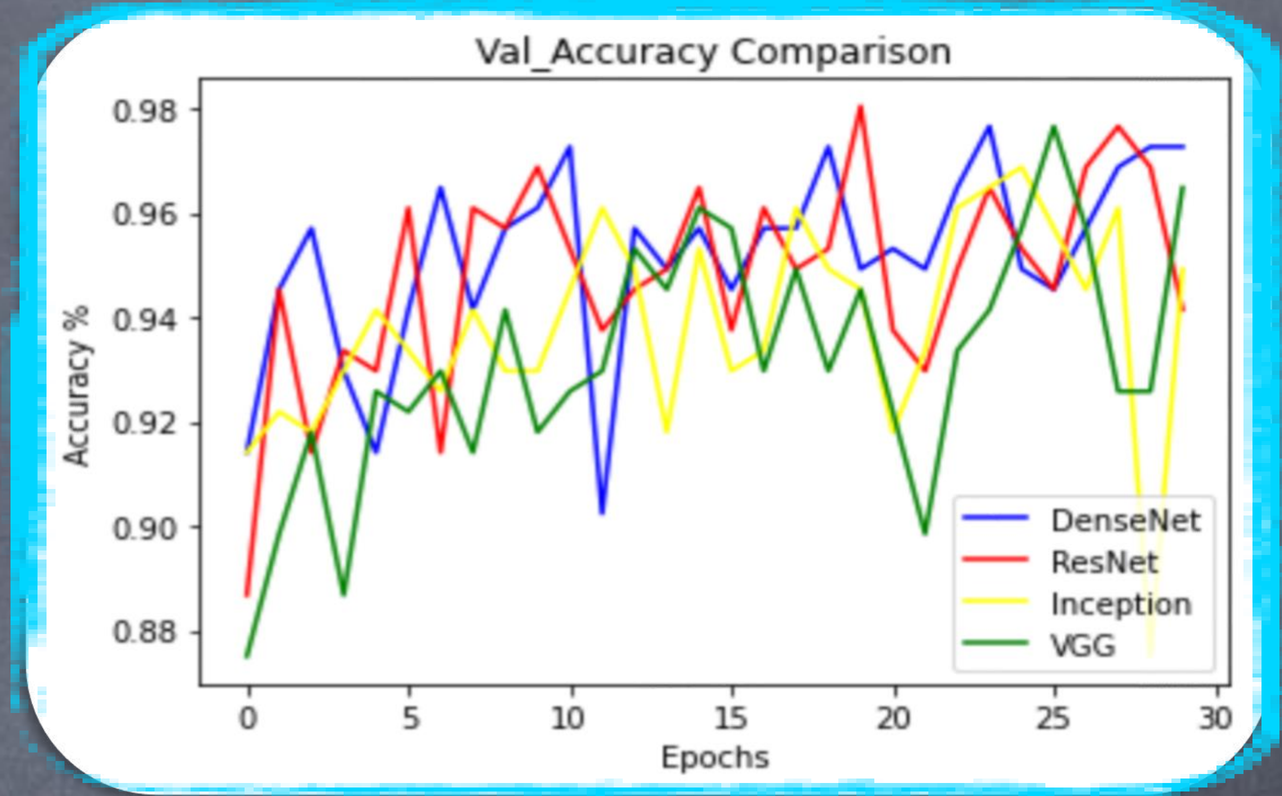
The CNN-based pre-trained models used to classify CXR images in this investigation is vgg16. Before finalizing the Model with VGG-16, a few other models were also trained and tested with the same dataset for a comparative study on the accuracy on each of them.

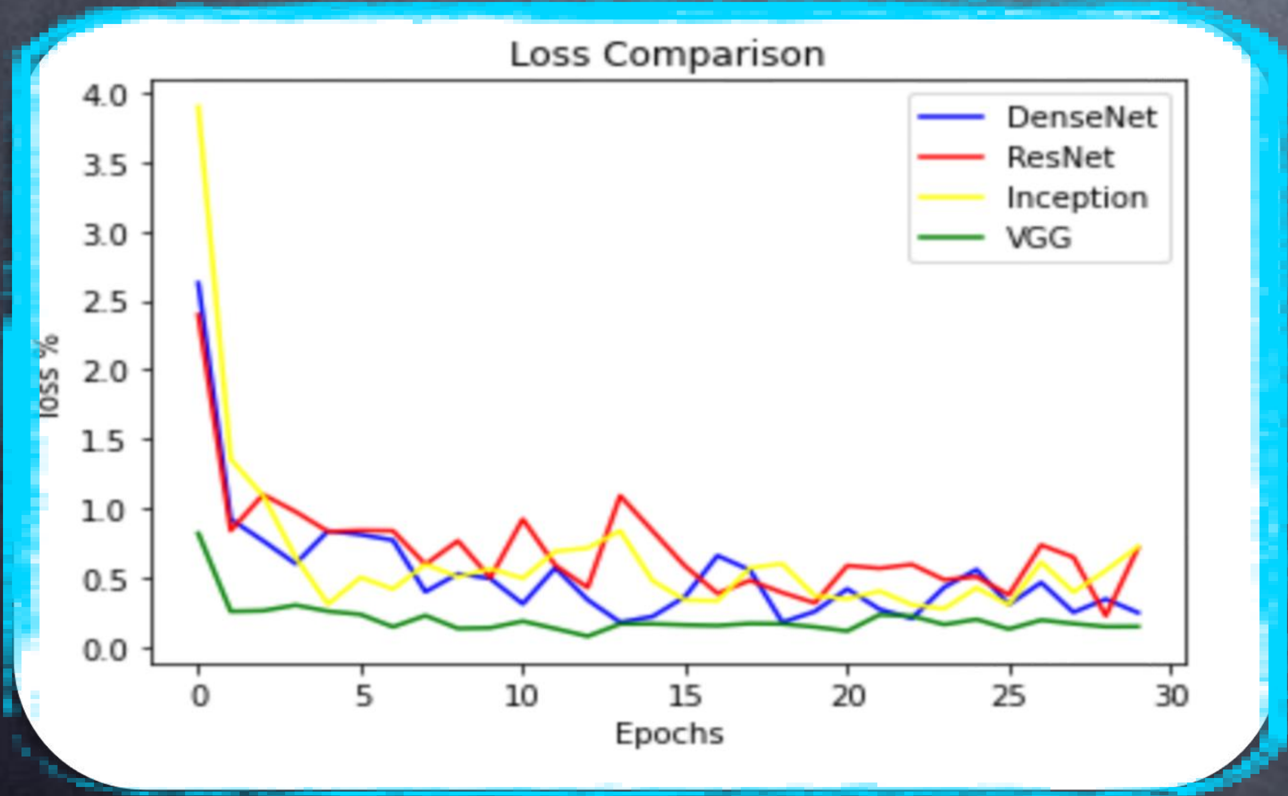
VGG16 is a convolution neural network (CNN) architecture that was awarded first place in the 2014 ILSVR (Imagenet) competition. It is widely recognised as one of the best vision model designs available today. Instead of a large number of hyper-parameters, VGG16 focuses on having convolution layers of 3x3 filter with stride 1 and always uses the same padding and maxpool layer of 2x2 filter with stride 2. This layout of convolution and max pool layers is maintained throughout the architecture. Finally, for output, it has two FC (fully connected layers) and a softmax. The number 16 in VGG16 alludes to the fact that it contains 16 weighted layers. This network is quite large, with an estimated 138 million parameters.

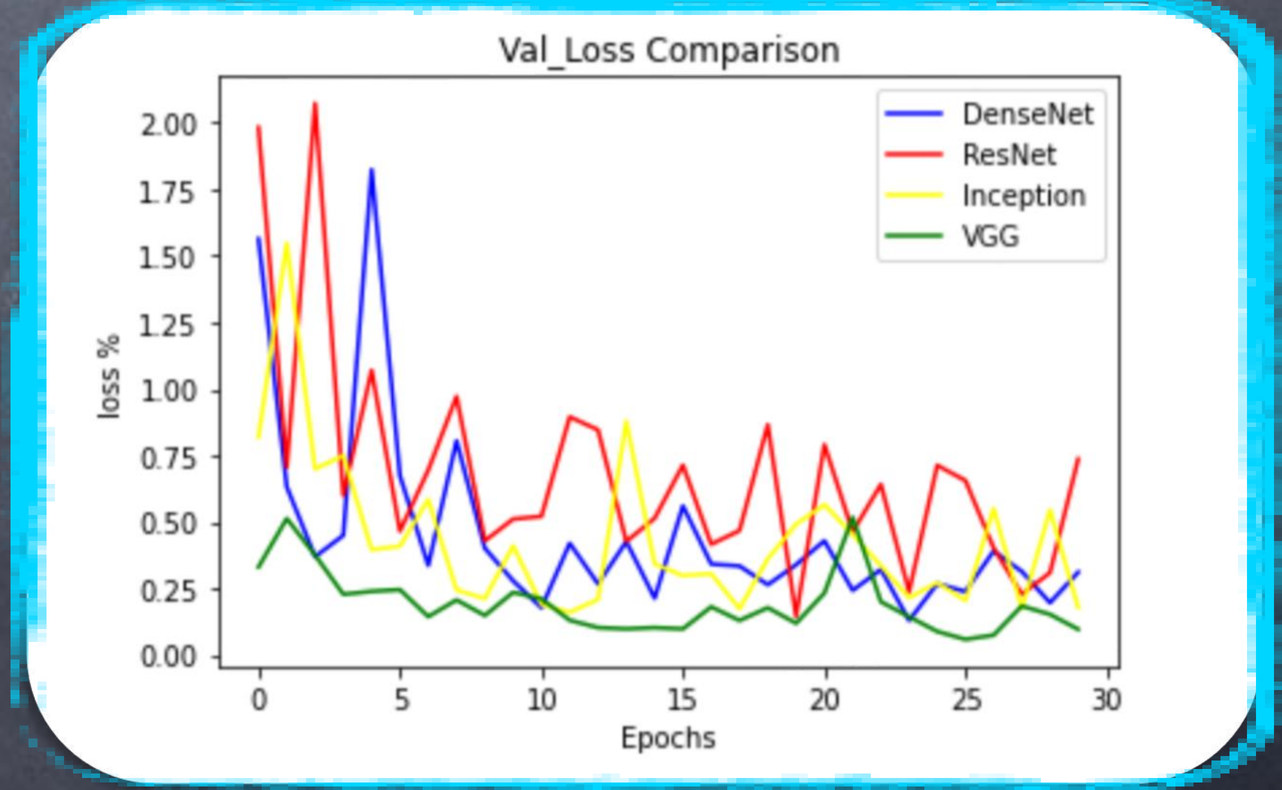


Model Comparison

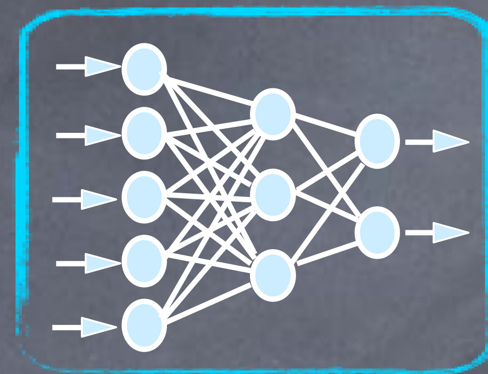








Model Test



- Covid Infected lungs
- Normal lungs

The given X-Ray image is of type = Covid

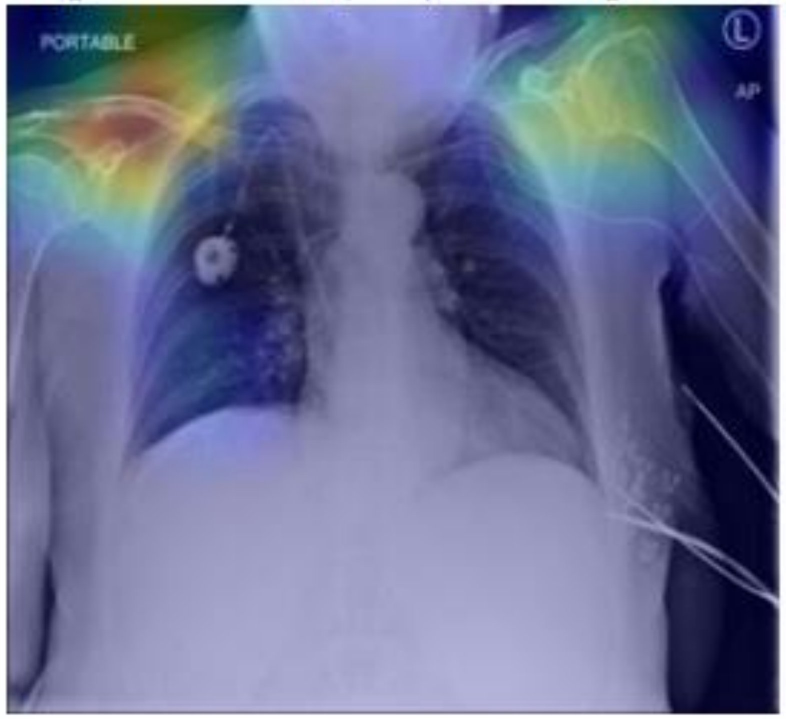
The chances of image being Covid is : 54.84541058540344 %
The chances of image being Normal is : 26.994240283966064 %

image with heatmap representing the covid spot

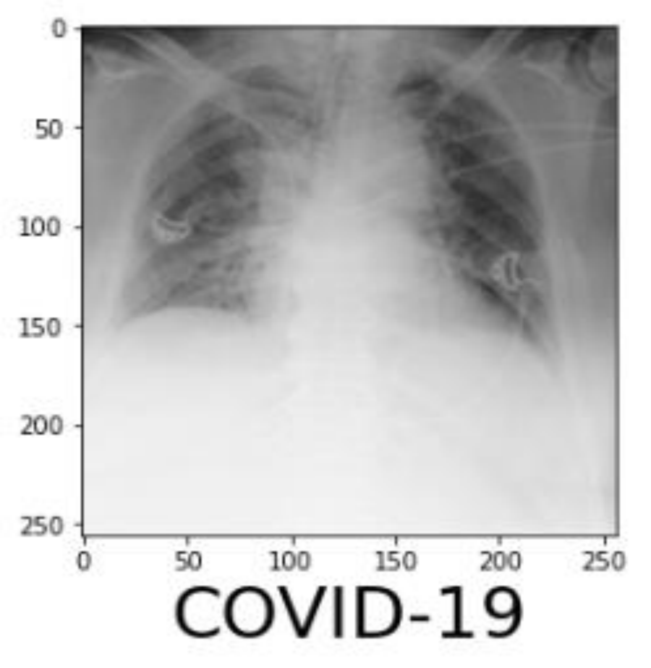
The given X-Ray image is of type = Normal

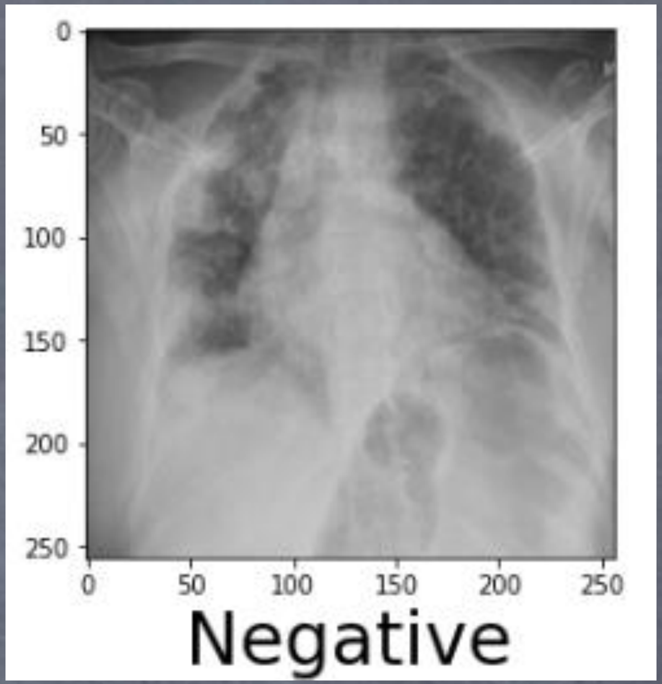
The chances of image being Covid is : 22.541259229183197 %
The chances of image being Normal is : 66.90725088119507 %

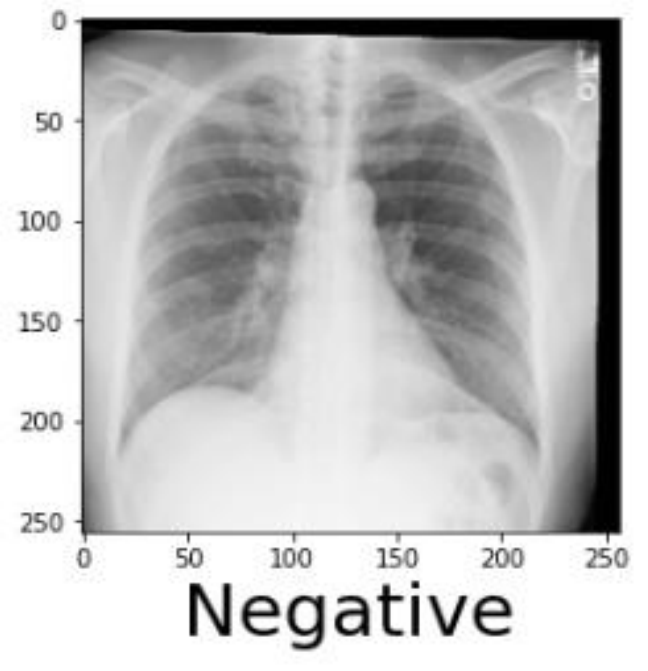
image with heatmap representing the covid spot

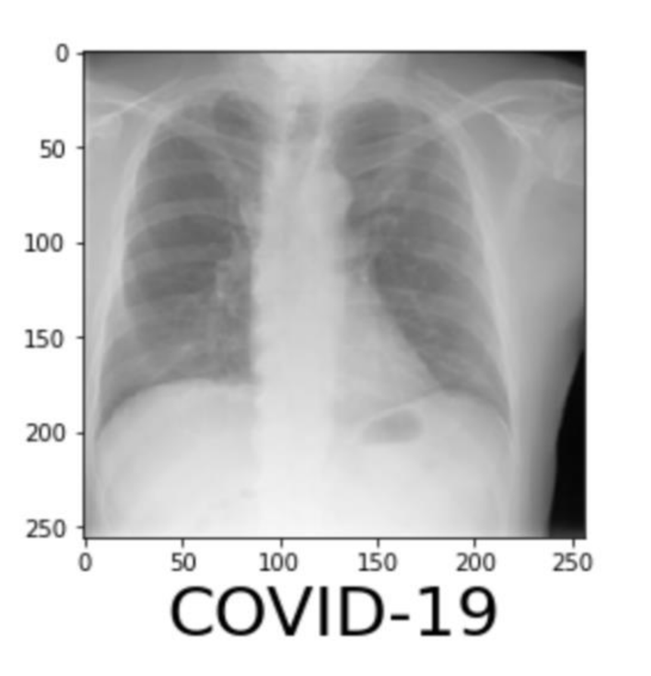


Outcomes

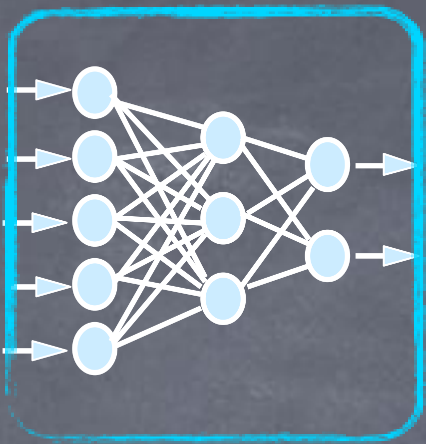

0
50
100
150
200
250
COVID-19


0
50
100
150
200
250
Negative


0
50
100
150
200
250
Negative


0
50
100
150
200
250
COVID-19

Conclusion



This innovation will greatly change the medical sector. Using this technique, COVID-19 patients can be quickly identified, which may contribute to addressing the current pandemic situation. Chest radiography is comparably safe for obtaining a sample than from the nose of a patient. In the future, this type of technique will help diagnosis. Several deep learning techniques can be used to optimise the parameters to create a robust model, which can help mankind. The metaheuristic-based deep COVID-19 model could also be a good technique to be explored in the future . Some more transfer learning-based models can be added in further development to compare the accuracy and optimisation of parameters , as well as a large dataset of normal and COVID-19 patients. The results can be observed by changing the ratio of training and testing data, and further comparative analysis can be performed.