Application of Deep Learning Image Classification Models to Detect Lung Diseases

Md Ayat Ullah, Md Sajid Hasan, Hossain Abdullah Taif

Advisor: Zunayeed Bin Zahir

Department of Electrical & Computer Engineering, North South University, Dhaka, Bangladesh

Abstract

Lung Diseases have become a severe adverse medical condition due to shoddy air quality, smoking, infections, genetic and other conditions. Early detection of lung diseases can prevent deaths of patients by taking cures. In our work, we aim to build an automated Computer Aided Diagnosis (CADx) system that can detect lung diseases similar to that of expert radiologists. Thus making x-ray diagnosis cheap and more available to the general population.

CNN-based classification model:

A neural network was created by mimicking how human brain process a task. Just like the neurons in brain communicating with each other through axon and dendrites our neural network communicate with each other through connecting path between them.

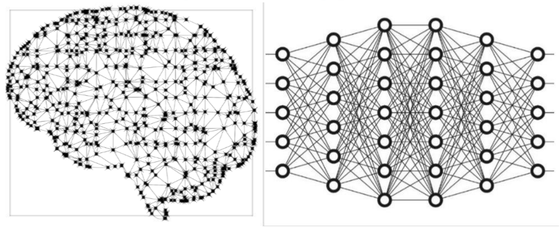


Figure: Artificial Neural Network inspired from human brain

Convolutional neural networks are inspired by how the visual cortex processes image. Features and patterns are extracted from images in convolution layer before sending them to a fully connected layer. Each neurons take inputs as weight, bias and activates for specific pattern through mathematical functions such as Relu (Rectified Linear Unit)

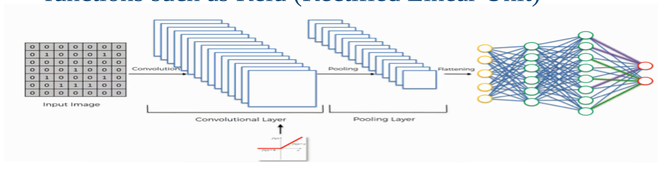


Figure: Convolution in Neural Network

Dataset

The National Insitute of Health (NIH) dataset comprises 112,120 frontal-view chest X-ray images of 30,805 unique patients with 14 disease labels. To evaluate the model, we randomly split the dataset into training (70%), validation (10%) and test (20%) sets, following the work in paper. Partitioned image names and corresponding labels are placed under the directory labels.

Flowchart and Architecture

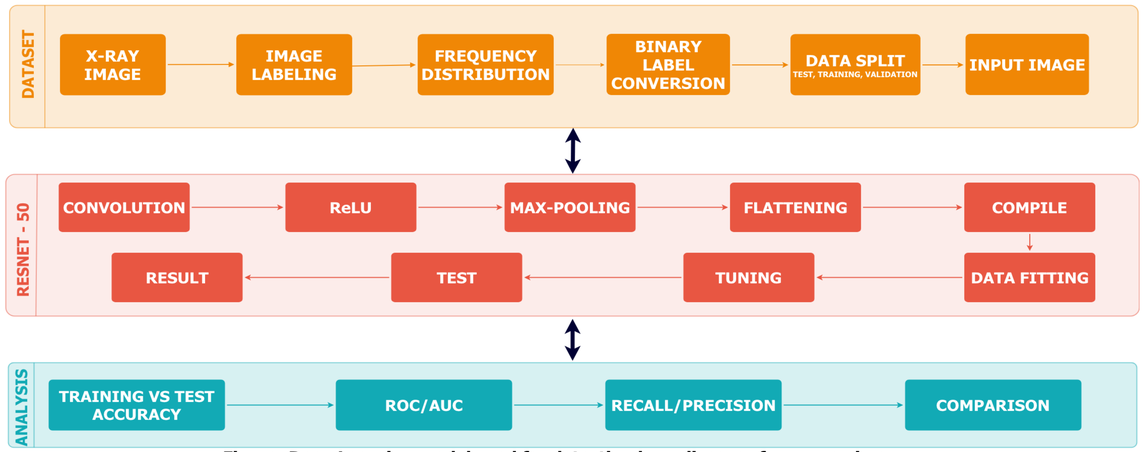


Figure: Deep Learning model used for detecting lung diseases from x-ray images.

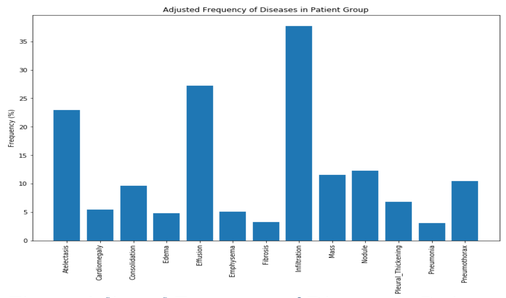


Figure: Adjusted Frequency of Diseases in Patients

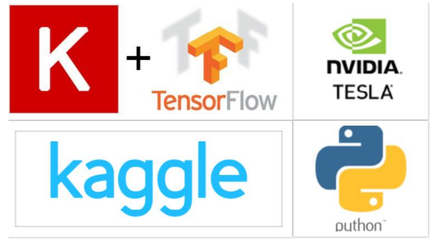
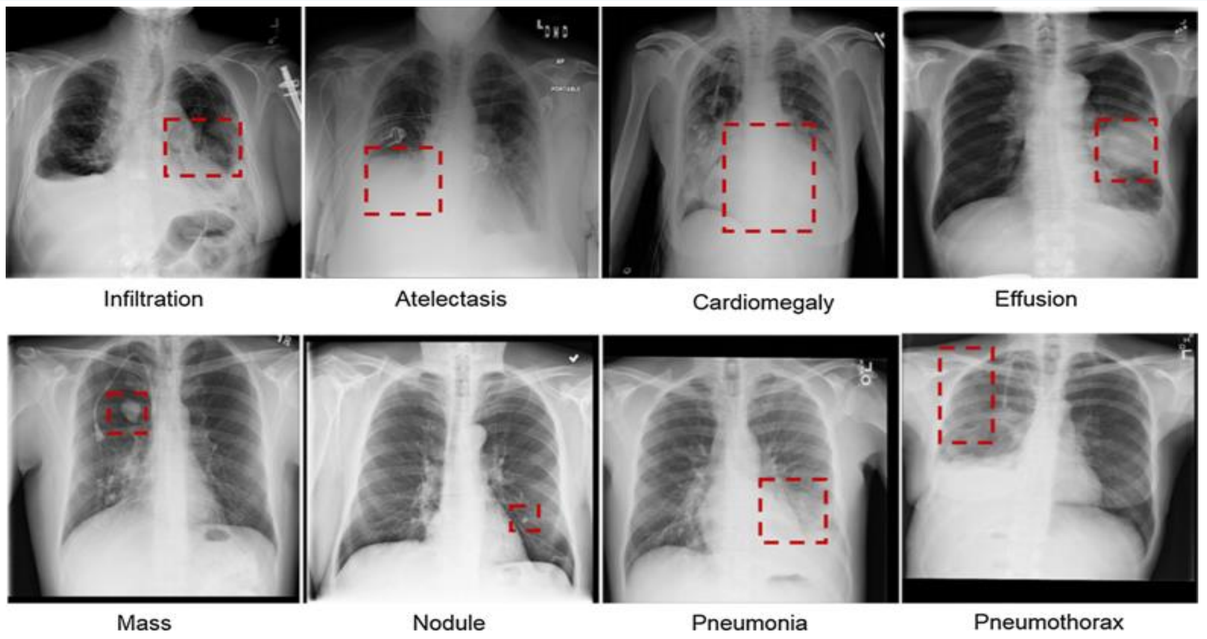


Figure: Platforms and Tools used in making the model

Dataset with identifying patterns



Result Analysis

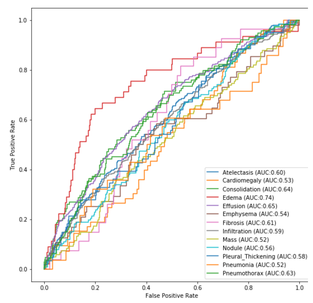


Figure: ROC Curve

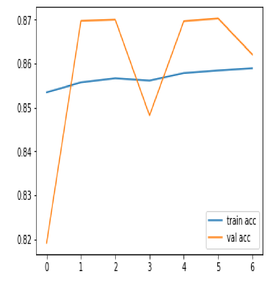


Figure: Test vs Validation accuracy

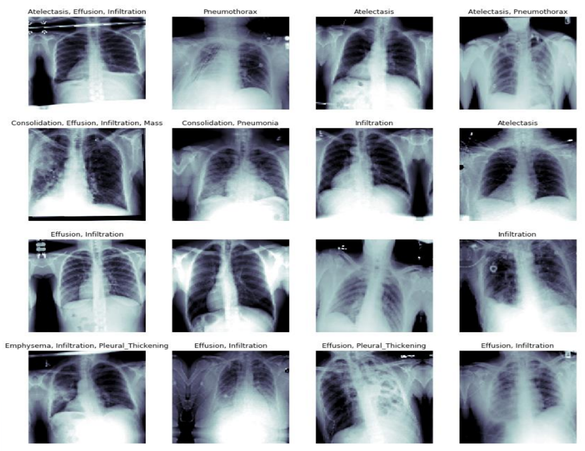


Figure: Images with associated predictions

Conclusion & Future Work

In our model we used NIH dataset which is a very sophisticated dataset for research purpose. To make our model ideal to be used in Bangladesh, local datasets are necessary due to mutation and variation across nations. Earlier we started to collect native dataset but it was postponed due to some difficulties. We will soon collect our local datasets and implement a model from that. At the same time we are working on the other models such as Mobilenet, Densenet and getting more than 85% validation accuracy. However we want to make an ensemble model which will give state of the art results and implement it in our local hospitals.