THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Machine Learning - UML501

COVID-19 DETECTION from CHEST X-RAY

Using Resnet-18 Convolutional Neural Network

By

Prajjwal Mehta | 102003127
Akshit KAthuria | 102003126

Submitted to

Niyaz Wani

Table of Contents

OBJECTIVE

To create an image classification model that can predict chest X-ray scans that belong to one of the three classes with a reasonably high accuracy. Please note that this dataset, and model that we train, can not be used to diagnose COVID-19 or Viral Pneumonia

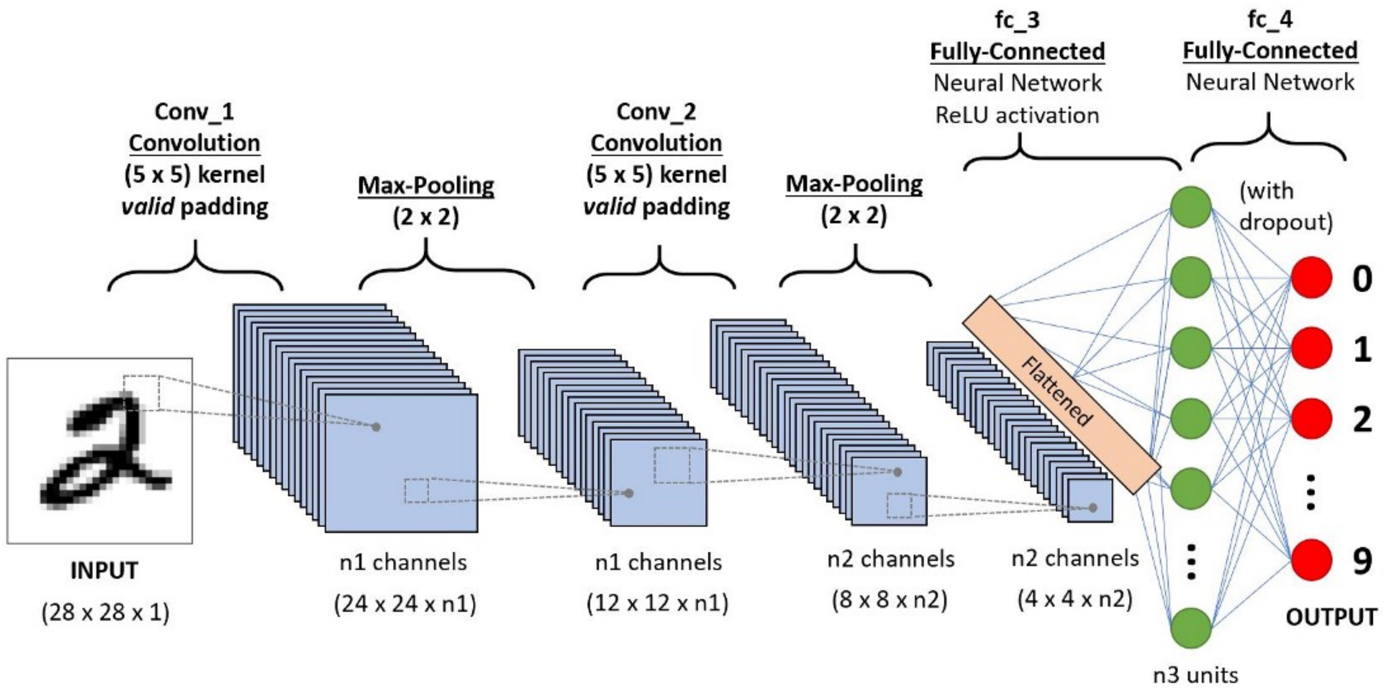
INTRODUCTION

X-rays are commonly used to diagnose various lung diseases, with the help of a trained medical professional. Through this project, we aim to aid medical professionals in the detection of lung diseases in the future. For the scope of this project, we have limited ourselves to the diagnosis of Pneumonia and Covid-19

We have used the Resnet-18 model from PyTorch which is a type of convolutional neural network that works to classify chest X-rays into one of three categories.

CONVOLUTIONAL NEURAL NETWORK

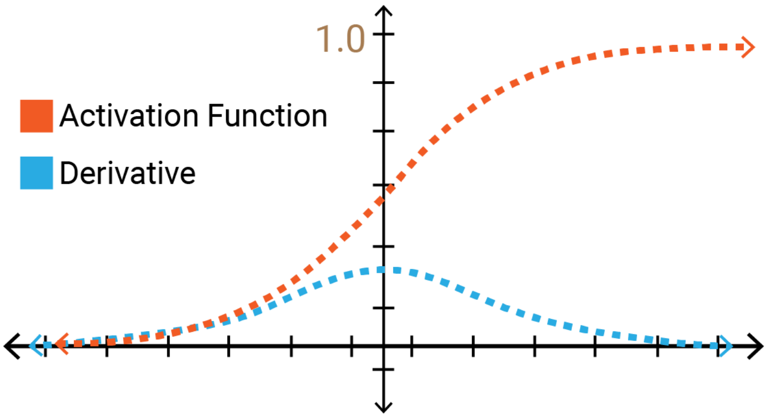
- A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.
- The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex



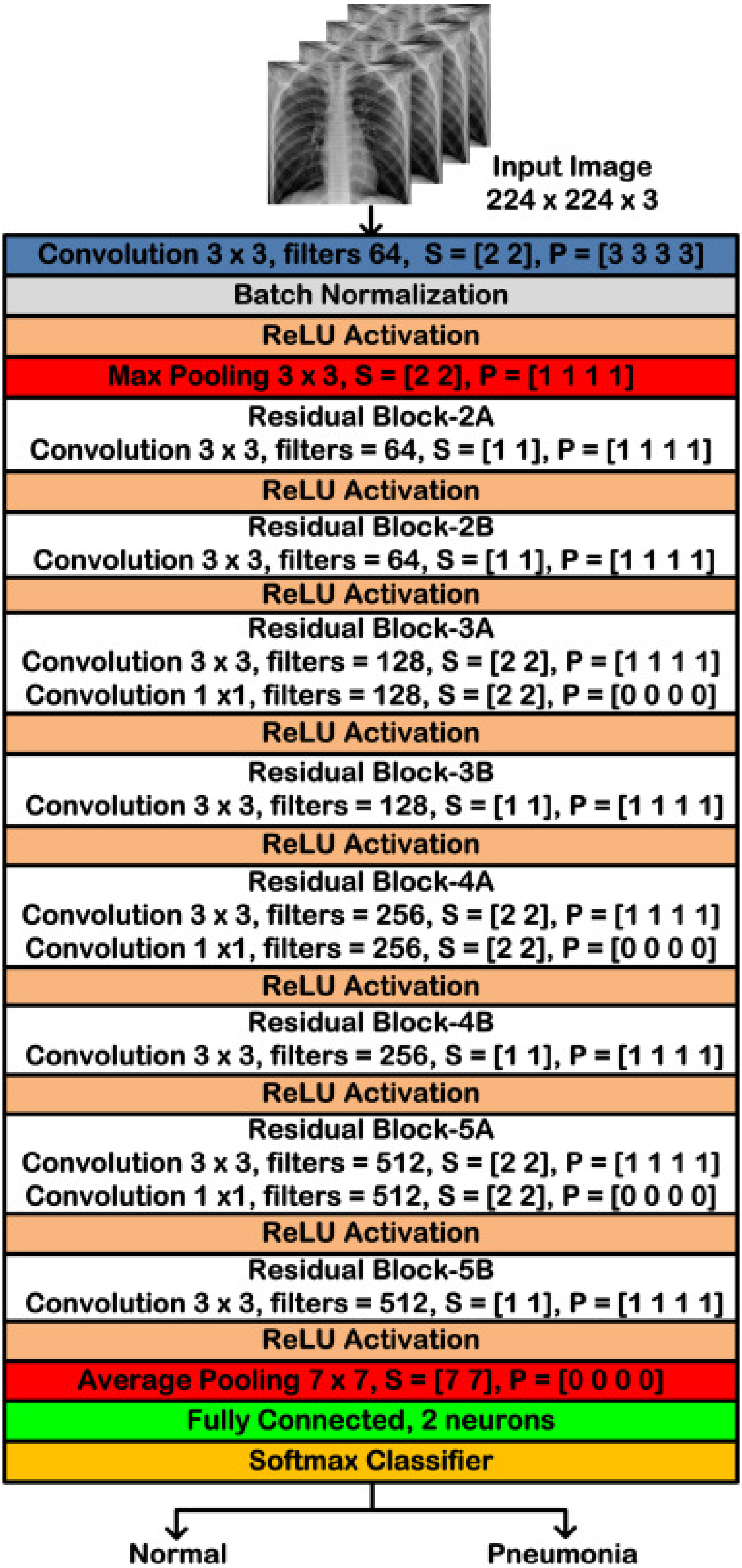
- An image is nothing but a matrix of pixel values, right? So why not just flatten the image (e.g. 3x3 image matrix into a 9x1 vector) and feed it to a Multi-Level Perceptron for classification purposes? Uh.. not really.In cases of extremely basic binary images, the method might show an average precision score while performing prediction of classes but would have little to no accuracy when it comes to complex images having pixel dependencies throughout.
- A ConvNet is able to successfully capture the Spatial and Temporal dependencies in an image through the application of relevant filters. The architecture performs a better fitting to the image dataset due to the reduction in the number of parameters involved and reusability of weights. In other words, the network can be trained to understand the sophistication of the image better.

RESNET-18

- The residual network has multiple variations, namely ResNet16, ResNet18, ResNet34, ResNet50, ResNet101, ResNet110, ResNet152, ResNet164, ResNet1202, and so forth.
- The ResNet stands for residual networks and was named by He et al. ResNet18 is a 72-layer architecture with 18 deep layers.
- The architecture of this network aimed at enabling large amounts of convolutional layers to function efficiently. However, the addition of multiple deep layers to a network often results in a degradation of the output.
- This is known as the problem of vanishing gradient where neural networks, while getting trained through back propagation, rely on the gradient descent, descending the loss function to find the minimizing weights.
- Due to the presence of multiple layers, the repeated multiplication results in the gradient becoming smaller and smaller thereby "vanishing" leading to a saturation in the network performance or even degrading the performance.



- The primary idea of ResNet is the use of jumping connections that are mostly referred to as shortcut connections or identity connections.
- These connections primarily function by hopping over one or multiple layers forming shortcuts between these layers.
- The aim of introducing these shortcut connections was to resolve the predominant issue of vanishing gradient faced by deep networks.
- These shortcut connections remove the vanishing gradient issue by again using the activations of the previous layer. These identity mappings initially do not do anything much except skip the connections, resulting in the use of previous layer activations.
- This process of skipping the connection compresses the network; hence, the network learns faster.
- This compression of the connections is followed by expansion of the layers so that the residual part of the network could also train and explore more feature space. The input size to the network is 224 × 224 × 3, which is predefined. The network is considered to be a DAG network due to its complex layered architecture and because the layers have input from multiple layers and give output to multiple layers.


Input Image
224 x 224 x 3
Convolution 3 x 3, filters 64, S = [2 2], P = [3 3 3 3]
Batch Normalization
ReLU Activation
Max Pooling 3 x 3, S = [2 2], P = [1 1 1 1]
Residual Block-2A
Convolution 3 x 3, filters = 64, S = [1 1], P = [1 1 1 1]
ReLU Activation
Residual Block-2B
Convolution 3 x 3, filters = 64, S = [1 1], P = [1 1 1 1]
ReLU Activation
Residual Block-3A
Convolution 3 x 3, filters = 128, S = [2 2], P = [1 1 1 1]
Convolution 1 x1, filters = 128, S = [2 2], P = [0 0 0 0]
ReLU Activation
Residual Block-3B
Convolution 3 x 3, filters = 128, S = [1 1], P = [1 1 1 1]
ReLU Activation
Residual Block-4A
Convolution 3 x 3, filters = 256, S = [2 2], P = [1 1 1 1]
Convolution 1 x1, filters = 256, S = [2 2], P = [0 0 0 0]
ReLU Activation
Residual Block-4B
Convolution 3 x 3, filters = 256, S = [1 1], P = [1 1 1 1]
ReLU Activation
Residual Block-5A
Convolution 3 x 3, filters = 512, S = [2 2], P = [1 1 1 1]
Convolution 1 x1, filters = 512, S = [2 2], P = [0 0 0 0]
ReLU Activation
Residual Block-5B
Convolution 3 x 3, filters = 512, S = [1 1], P = [1 1 1 1]
ReLU Activation
Average Pooling 7 x 7, S = [7 7], P = [0 0 0 0]
Fully Connected, 2 neurons
Softmax Classifier
Normal
Pneumonia

DATASET DESCRIPTION

The dataset used for this project is [COVID-19 Radiography Dataset] available on Kaggle. The dataset consists of 33,920 X-rays which are divided into the following categories

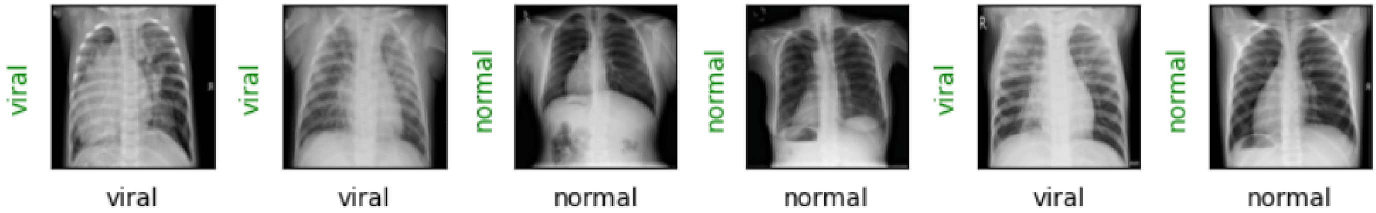
Categories :

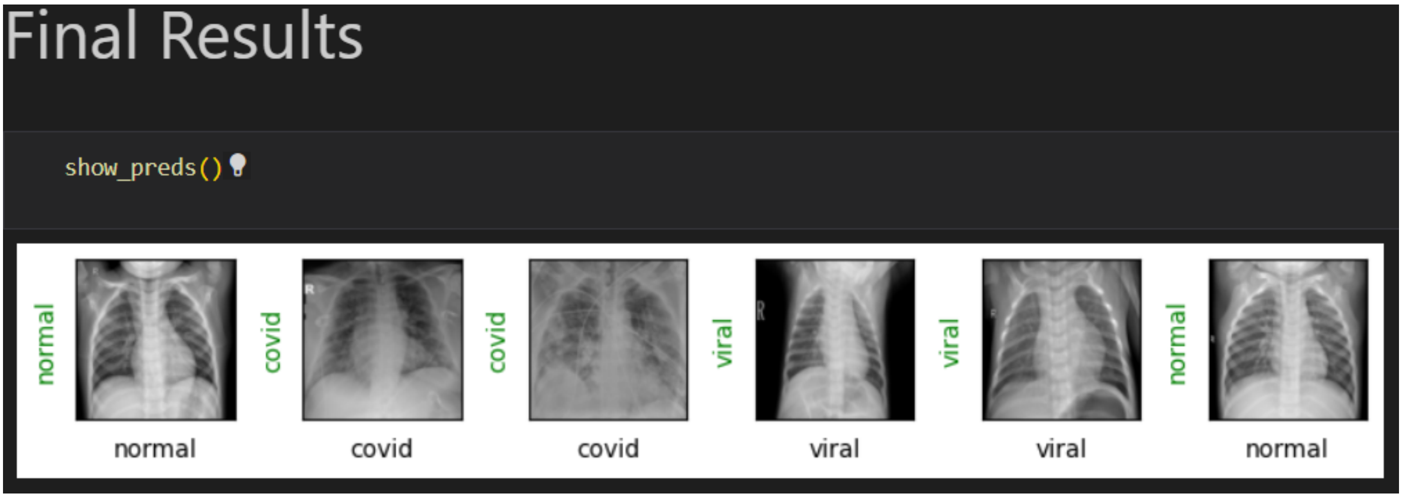
- Viral Pneumonia
  - Total images : 1345
  - Images used : 1345
- Covid-19
  - Total images : 3616
  - Images used : 1604
- Normal
  - Total images : 10192
  - Images used : 1447

Further, in order to get the most out of our dataset, we have randomly flipped some images in the training dataset.

RESULTS

- Using our trained resnet-18 model, we have labeled the x-rays as either normal, viral or covid.





FUTURE SCOPE

- The model can also be trained to work for colored X-rays
- The model can be used for preliminary diagnosis of lung diseases