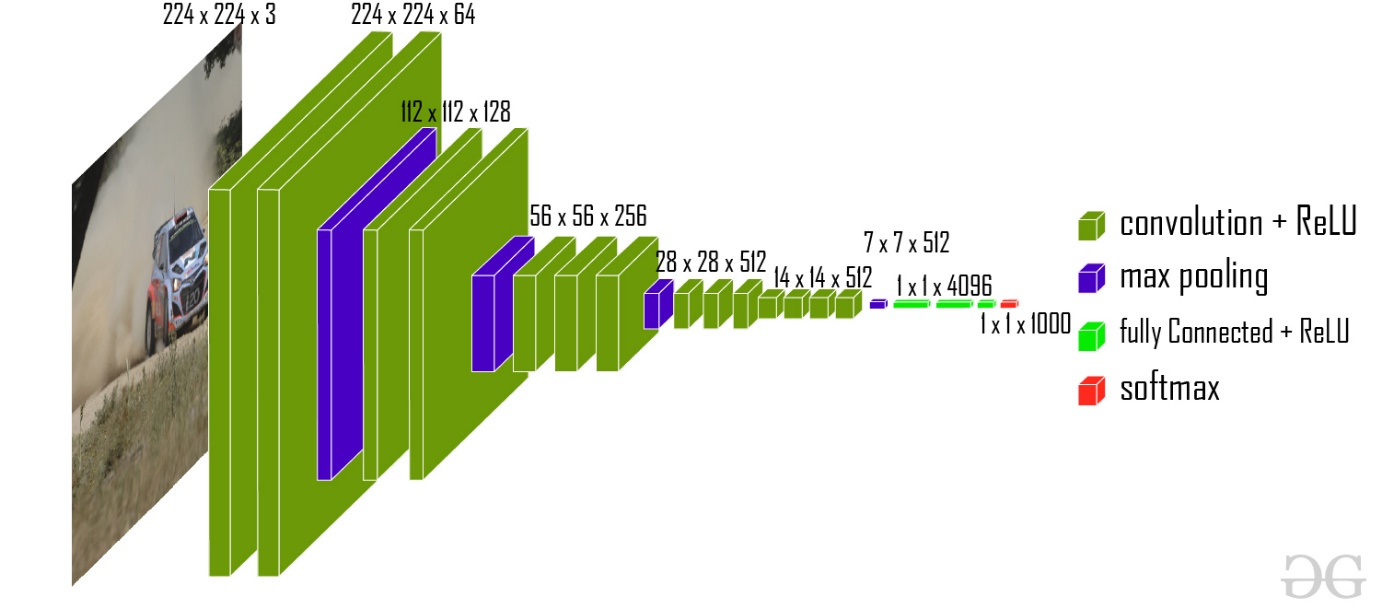
**Tomato Image Detection and Classification:**

**Agriculture is the backbone of India Economy. Still more than 70% people dpend on it.**

Plant diseases put on a heavy toll on the agricultural economy. Early detection and prevention project promising results in cutting losses. We propose an software solution to automatically detect disease and classify which type of disease. We are using VGG model for classification of Image. In the VGG model we have used only 8 Layers for our model.

VGG 16 was proposed by Karen Simonyan and Andrew Zisserman of the Visual Geometry Group Lab of Oxford University in 2014 in the paper “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION”.

The input to the network is image of dimensions (224, 224, 3). The first two layers have 64 channels of 3\*3 filter size and same padding. Then after a max pool layer of stride (2, 2), two layers which have convolution layers of 256 filter size and filter size (3, 3). This followed by a max pooling layer of stride (2, 2) which is same as previous layer. Then there are 2 convolution layers of filter size (3, 3) and 256 filter. After that there are 2 sets of 3 convolution layer and a max pool layer. Each have 512 filters of (3, 3) size with same padding.This image is then passed to the stack of two convolution layers. In these convolution and max pooling layers, the filters we use is of the size 3\*3 instead of 11\*11 in AlexNet and 7\*7 in ZF-Net. In some of the layers, it also uses 1\*1 pixel which is used to manipulate the number of input channels. There is a padding of 1-pixel (same padding) done after each convolution layer to prevent the spatial feature of the image.

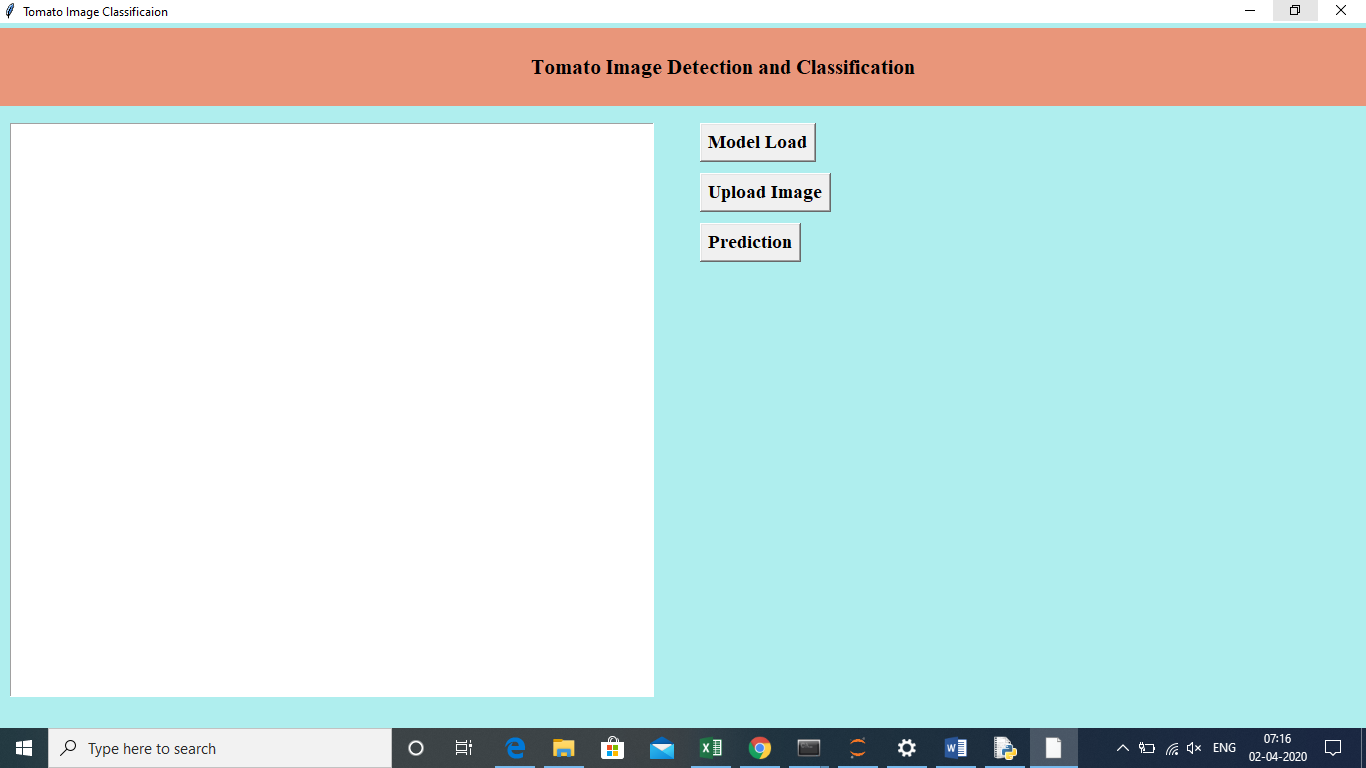
****

**Architecture:**   
The input to the network is image of dimensions (224, 224, 3). The first two layers have 64 channels of 3\*3 filter size and same padding. Then after a max pool layer of stride (2, 2), two layers which have convolution layers of 256 filter size and filter size (3, 3). This followed by a max pooling layer of stride (2, 2) which is same as previous layer. Then there are 2 convolution layers of filter size (3, 3) and 256 filter. After that there are 2 sets of 3 convolution layer and a max pool layer. Each have 512 filters of (3, 3) size with same padding.This image is then passed to the stack of two convolution layers. In these convolution and max pooling layers, the filters we use is of the size 3\*3 instead of 11\*11 in AlexNet and 7\*7 in ZF-Net. In some of the layers, it also uses 1\*1 pixel which is used to manipulate the number of input channels. There is a padding of 1-pixel (same padding) done after each convolution layer to prevent the spatial feature of the image.

After the stack of convolution and max-pooling layer, we got a (7, 7, 512) feature map. We flatten this output to make it a (1, 25088) feature vector.After this there are 3 fully connected layer, the first layer takes input from the last feature vector and outputs a (1, 4096) vector, second layer also outputs a vector of size (1, 4096) but the third layer output a 1000 channels for 1000 classes of ILSVRC challenge, then after the output of 3rd fully connected layer is passed to softmax layer in order to normalize the classification vector. After the output of classification vector top-5 categories for evaluation. All the hidden layers use ReLU as its activation function. ReLU is more computationally efficient because it results in faster learning and it also decreases the likelihood of vanishing gradient problem..

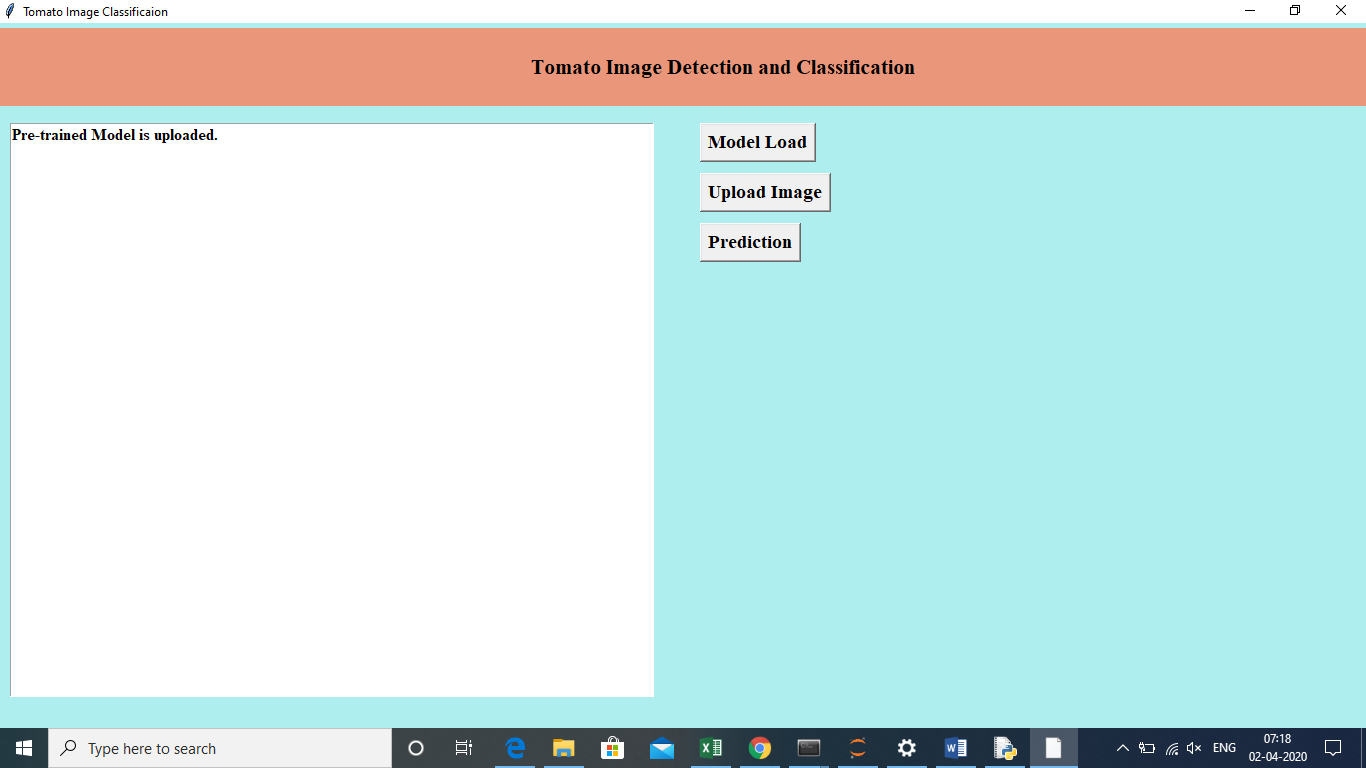
Steps for Execution:

1. Open Annaconda Prompt:Goto Project Location:
2. Run Python final-predict.py

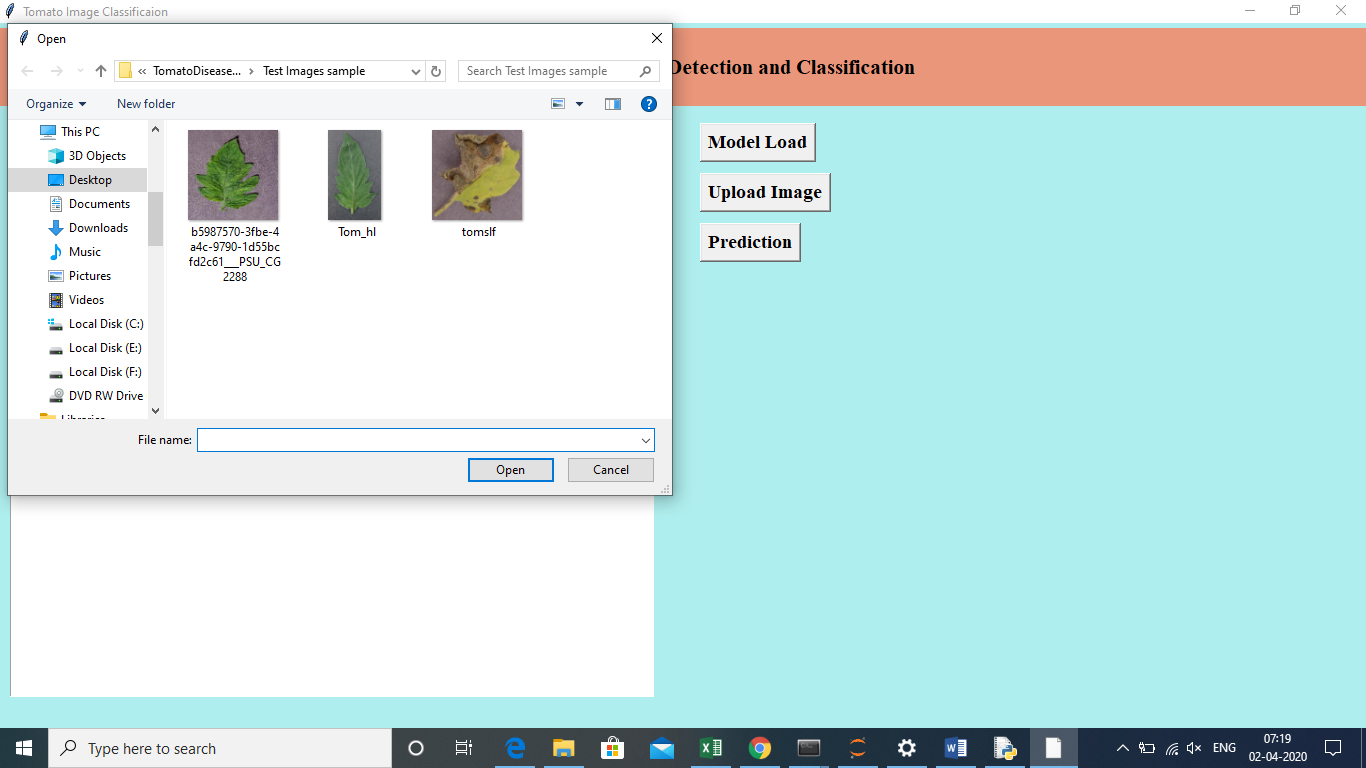


Will open the Desktop application.

1. Click on Model Load which will load Pre-trained module.



1. Then click on the Upload Image.



Upload the test Image to test on the model.

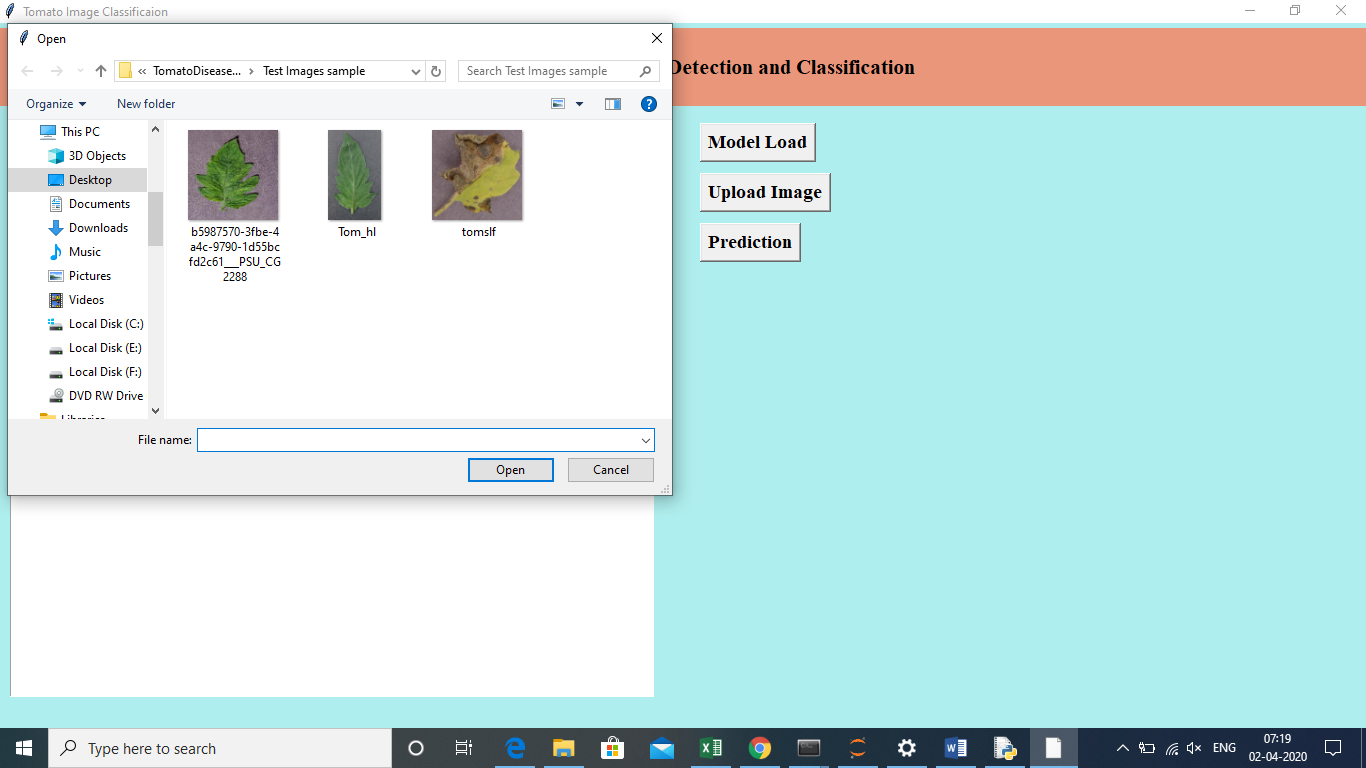
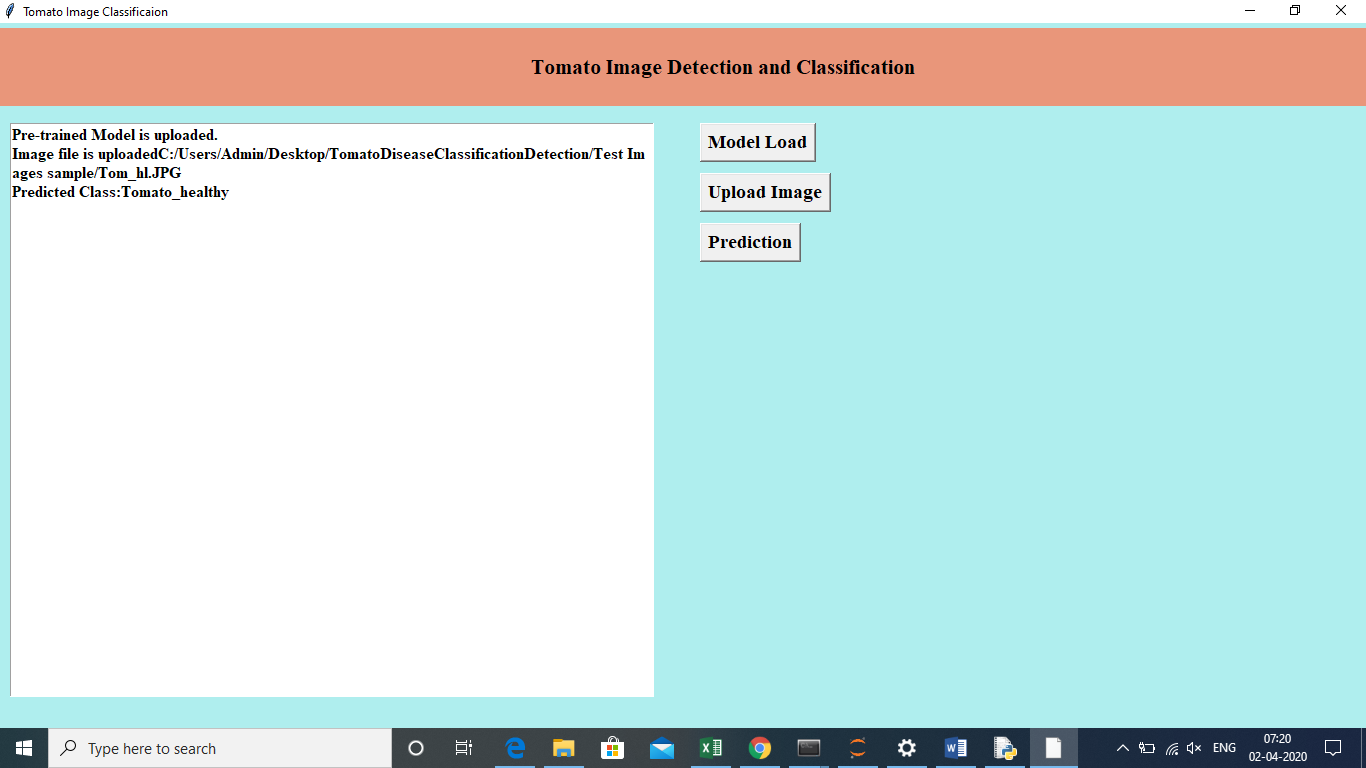


Image uploaded and Now we can test the Image.

1. Now click on the Prediction.



For the uploaded image it showing class as “Tomato Healthy”.

Note: For some images it won’t work proplerly. We have to train the model with more layers and more epoch operations.