CMPT412 Project 1 Report

Ritika Goyal (301401516)

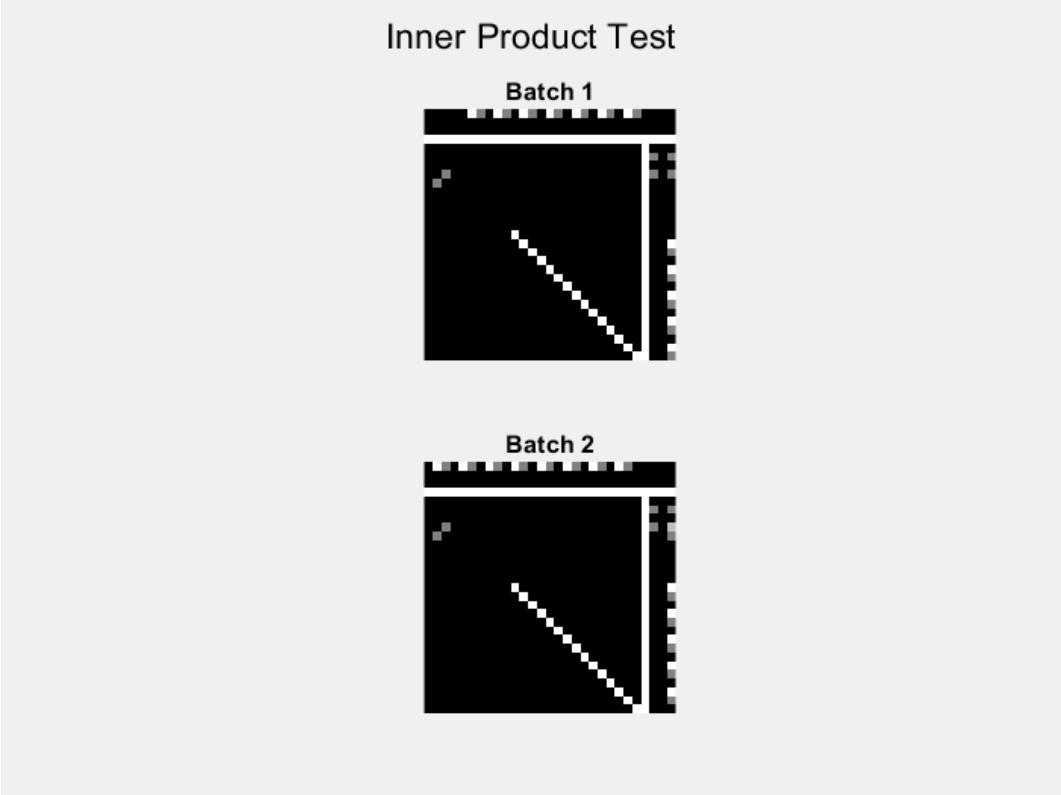
Part 1: Forward Pass

Q 1.1 Inner Product Layer:

In this layer, the output is calculated using Wx + b, where W is weight matrix, x is input, and b is bias.

First the outputs were specified and then all the images were traversed in a batch and output was calculated.

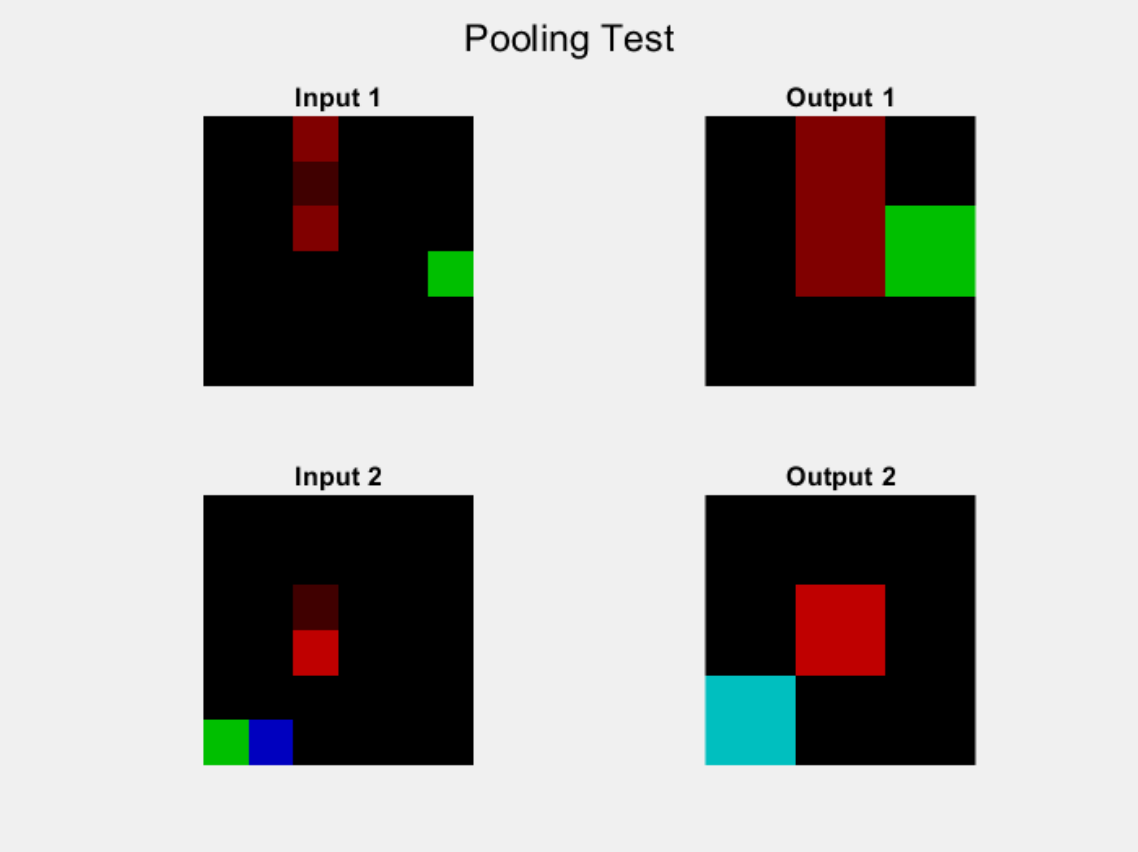
When test_components.m was executed, the inner product was visualized by following image.



Q 1.2: Pooling Layer:

In the pooling layer, the max pooling was used which reduces the size of feature maps by finding max value in each kernel.

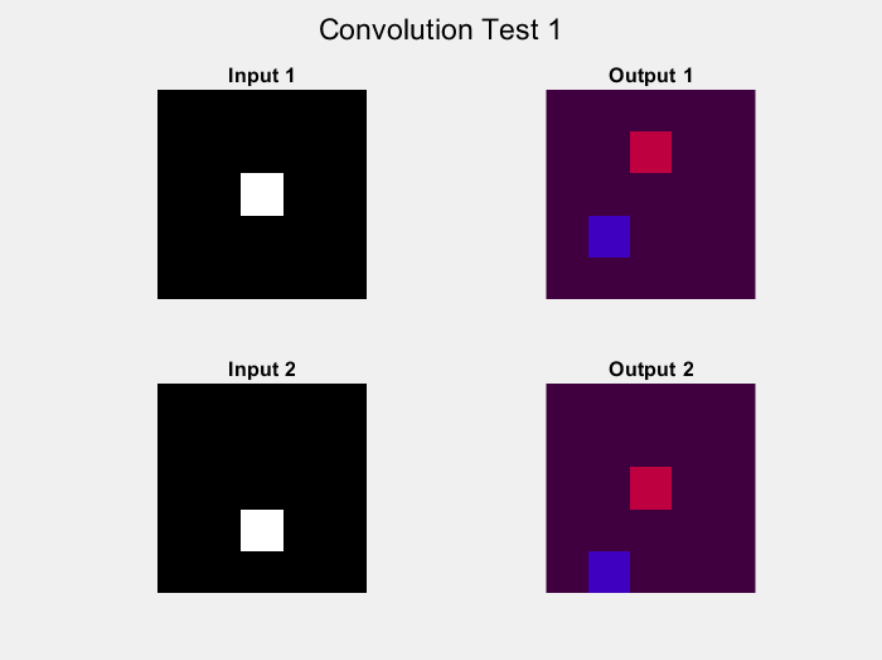
When test_components.m was executed, the polling was visualized by following image.

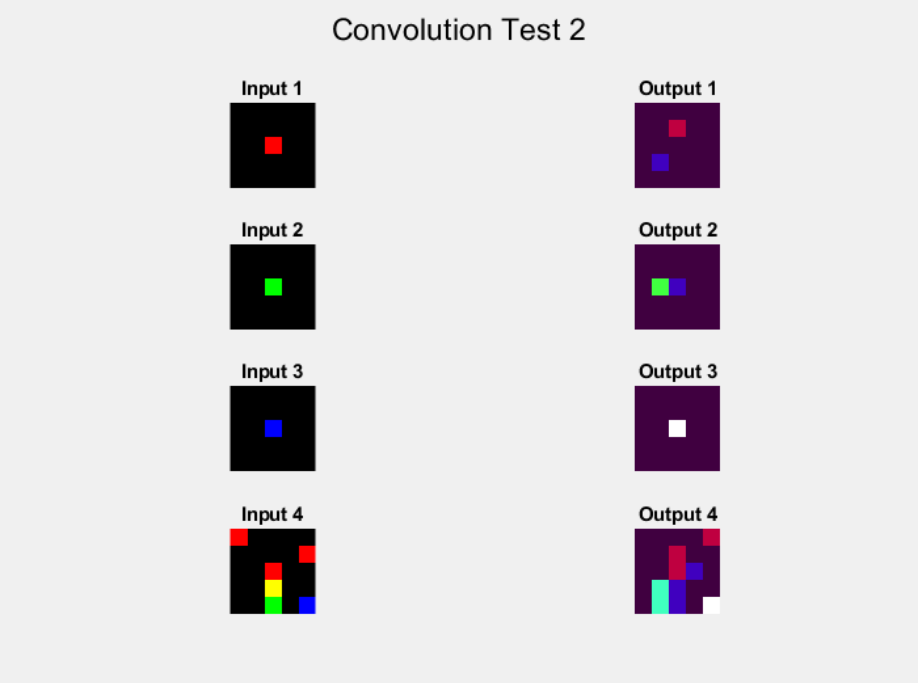


Q 1.3 Convolution Layer:

Convolution is performed using k X k filters and a W X H image. After convolution, feature map is obtained. Convolution layers reduces parameters and exploits spatial structure.

When test_components.m was executed, the convolution was visualized by following image.





Q 1.4 ReLU:

In Relu, f(x) = max (x,0) was used. If input is greater than or equal to 0, then output will be x (equal to input) but if input is less than 0, then the output will be 0 after this layer.

PART 2: BACK PROPAGATION

Q 2.1 ReLU:

The negative values are replaced by 0 and positive values are kept in matrix which are the multiplied element wise with output and result is stored in input_od.

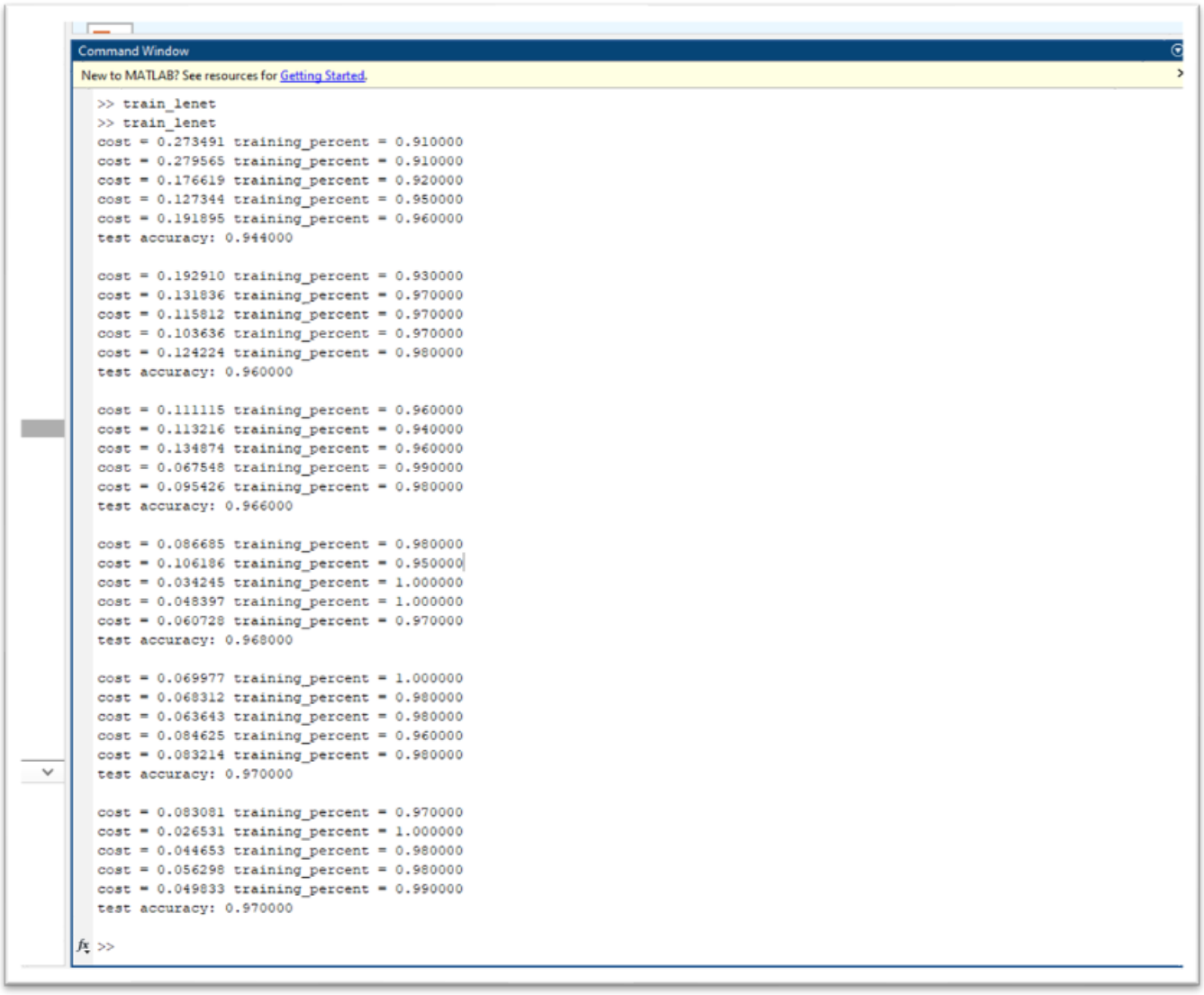
Q 2.2 Inner :Product Layer:

The differentiation output.diff is calculated.

PART 3: TRAINING

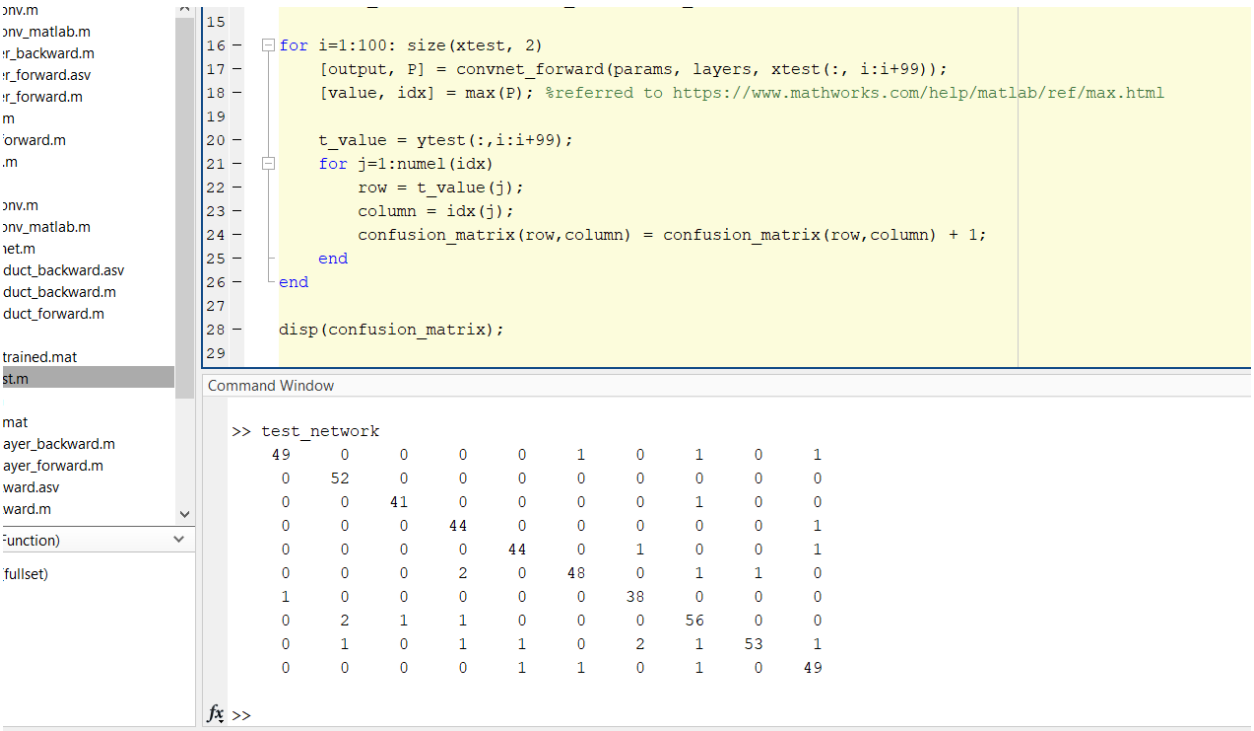
Q 3.1: Training:

After running train_lenet.m, the test accuracy of the CNN network is 97%.



Q 3.2: TEST THE NETWORK

When test_network.m was executed, the following confusion matrix was obtained:



It is obvious that if each time the test_network.m will be executed, the confusion_matrix would be different. According to above confusion matrix, the two most confused pairs were (2,7) and (4,9). Since these numbers can be easily mistaken by each other, so this confusion is justified.

Q 3.3 REAL WORLD TESTING

I downloaded 5 images and saved them in images/real_images. These images were passed through the network and some of them were correctly predicted.

Output -> 3

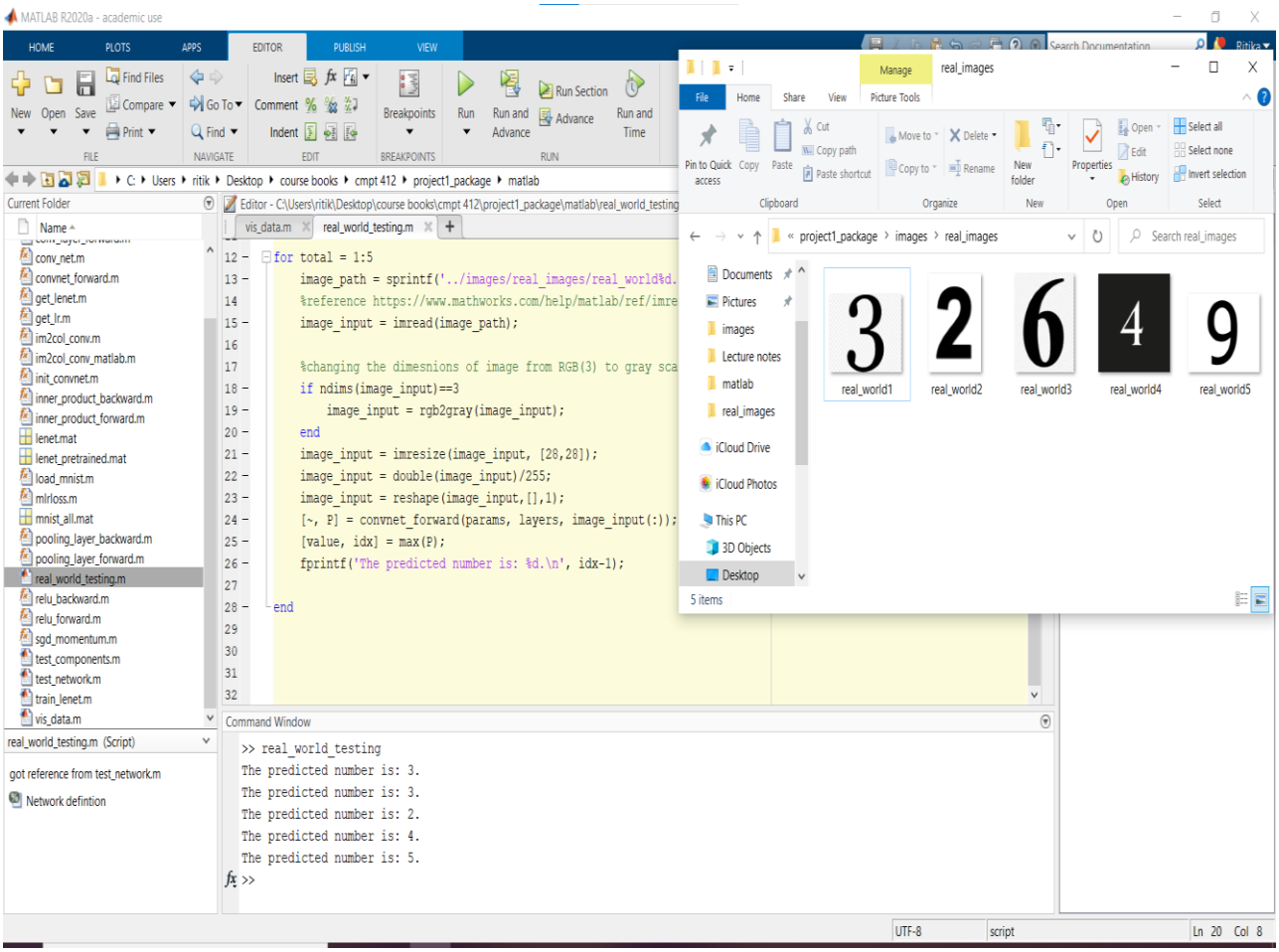
Output -> 3

Output -> 2

Output ->4

9

Output->5

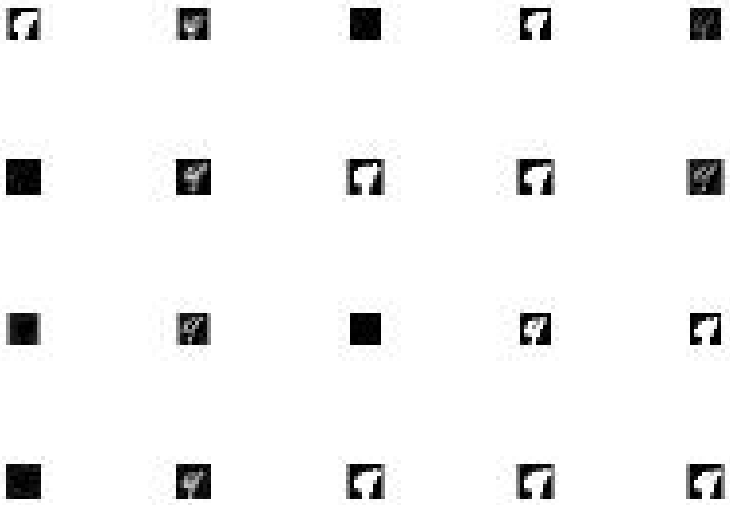

real_world_testing
The predicted number is: 3.
The predicted number is: 3.
The predicted number is: 2.
The predicted number is: 4.
The predicted number is: 5.
real_world1
real_world2
real_world3
real_world4
real_world5

Part 4: VISUALIZATION

Q 4.1:

After running vis_data.m, following outputs were obtained for convolution and Relu Layer.

CONV Layer:



ReLU Layer:

Q 4.2:

The layer filters are applied for the original image which is shown in feature map. The original image was digit 9. After passing the original image from both the layers, it is seen that both the layers are different. The size os the image changes from Conv to Relu. Relu adds more darkness to the image because Relu focuses to increase the non-linearity of images.

PART 5: IMAGE CLASSIFICATION

First the image is converted to greyscale by using threshold function. Then the different connected components were calculated and bounding box was assigned to components. After fixing the padding and resizing it, the image was passed through the network.