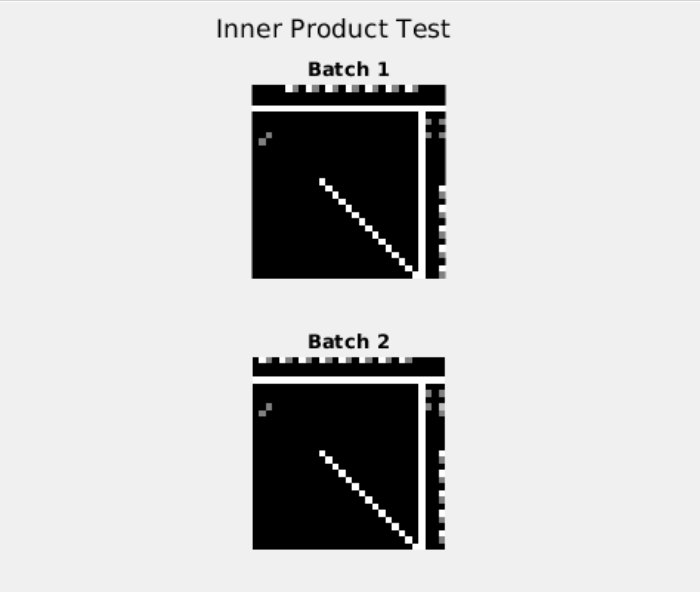
Darryl Julius Basri
301388030

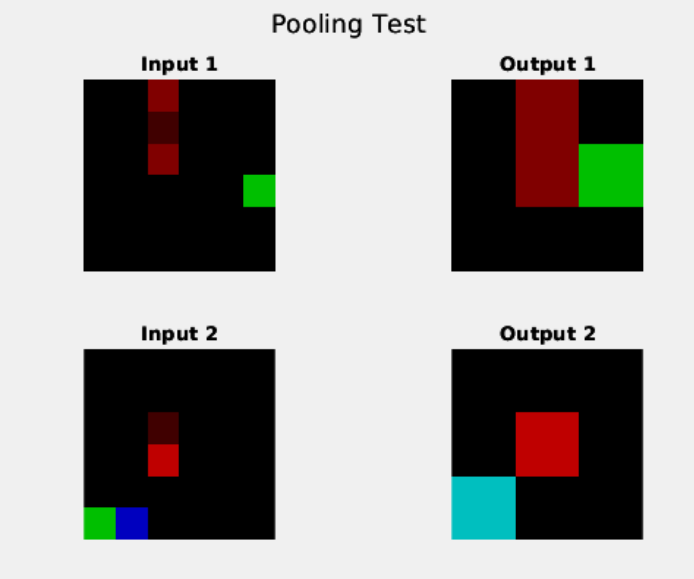
Free days used : 4

Part 1

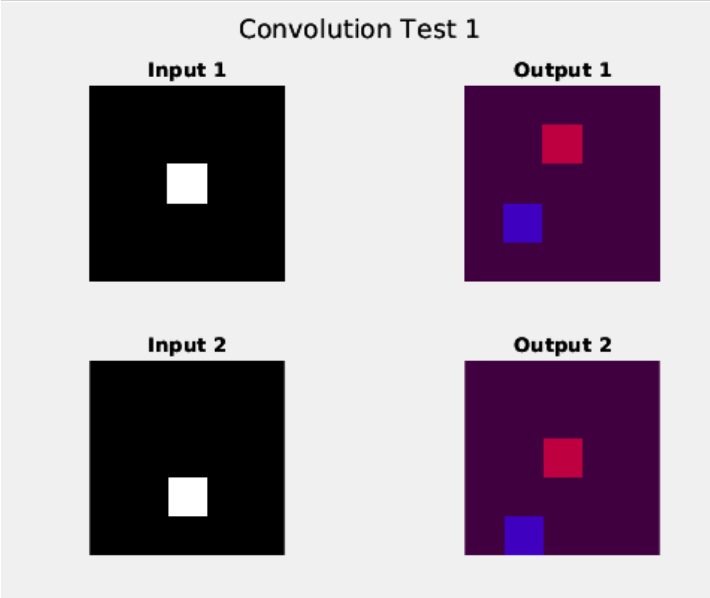
1.1



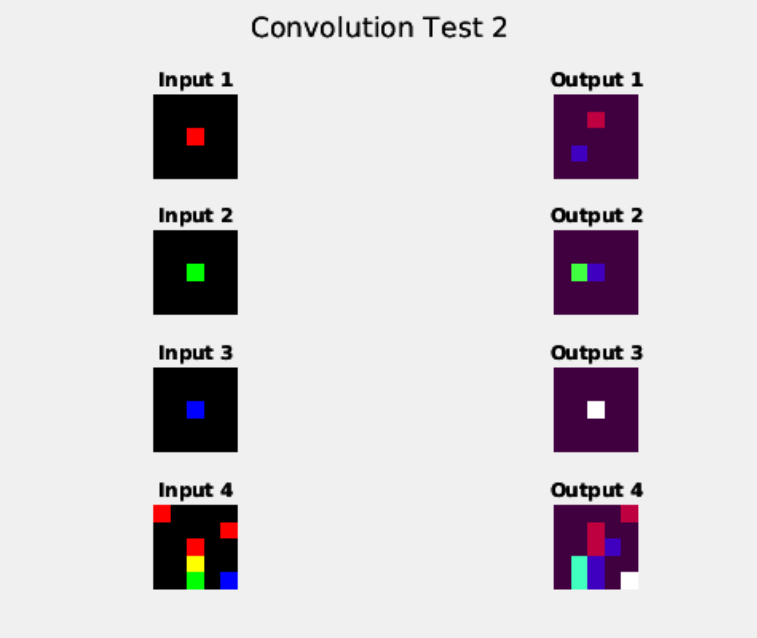
1.2



1.3


Convolution Test 1
Input 1
Output 1
Input 2
Output 2

1.4


Convolution Test 2
Input 1
Output 1
Input 2
Output 2
Input 3
Output 3
Input 4
Output 4

Part 3

3.1

```
cost = 0.273491 training_percent = 0.910000
cost = 0.279565 training_percent = 0.910000
cost = 0.176619 training_percent = 0.920000
cost = 0.127344 training_percent = 0.950000
cost = 0.191895 training_percent = 0.960000
test accuracy: 0.944000

cost = 0.192910 training_percent = 0.930000
cost = 0.131836 training_percent = 0.970000
cost = 0.115812 training_percent = 0.970000
cost = 0.103636 training_percent = 0.970000
cost = 0.124224 training_percent = 0.980000
test accuracy: 0.960000
```

3.2

```
confusion_matrix =

     8     0     0     0     0     0     0     0     1     0
     0    17     0     0     0     0     0     0     0     0
     0     0     7     0     1     0     0     1     0     0
     0     0     0     7     0     0     0     0     0     0
     0     0     0     0    13     0     0     0     0     0
     0     0     0     0     0     8     0     0     0     0
     0     0     0     0     0     0    11     0     0     0
     0     0     0     0     0     0     0     7     0     0
     0     0     0     0     0     0     0     1     6     0
     0     0     0     0     0     0     0     1     0    11
```

From the confusion matrix above, we can see that the top two most confusing class is the number 8, when pairs with the numbers 3, 9 ,0. Which makes sense because the number 8 is very similar to the number 3 in its key components which can be confused by the network.

3.3

```
>> testing_rw
predicted_labels =
     5
```

```
>> testing_rw
predicted_labels =
     1
```

```
>> testing_rw
predicted_labels =
     7
```

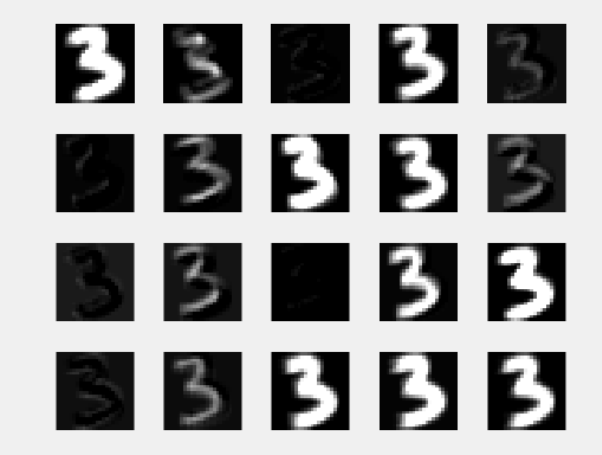
```
>> testing_rw
predicted_labels =
     3
```

```
>> testing_rw

predicted_labels =

     3
```

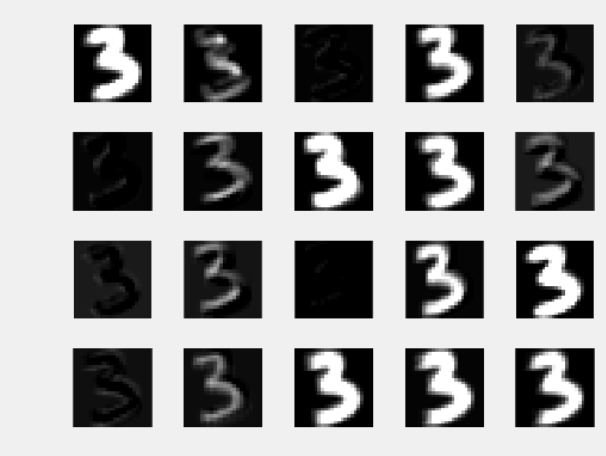
Part 4

4.1

Original Image



CONV Layer



ReLu Layer

4.2

The features image from the convolution layer is a filtered version of the original image, where the filter (or kernel) is used to highlight certain patterns or features in the image. The features image from the ReLU layer would be a version of the filtered image where any negative values are set to zero. This result in a sparser feature representation, as well as reducing the chance of vanishing gradients.

Part 5

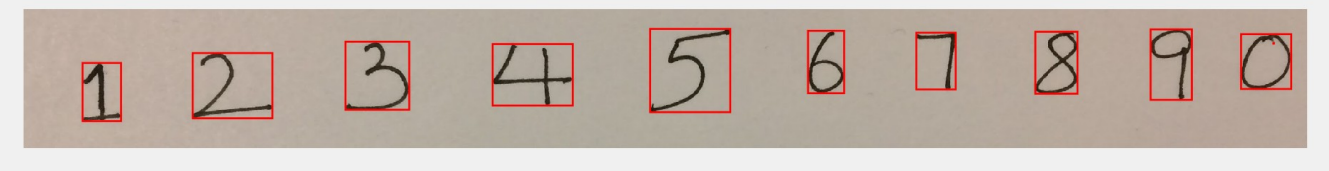


Image 1

For Image 1, the network predicted : (in order from left to right)

```
prediction =

     8     5     6     3     5     8     3     5     5     3     8
```

As can be seen above, the network produced 11 results when there are only 4 digits. This is because the binding box pick up an extra components on the 0 digits. The resulting accuracy from Image1 is 1/10.

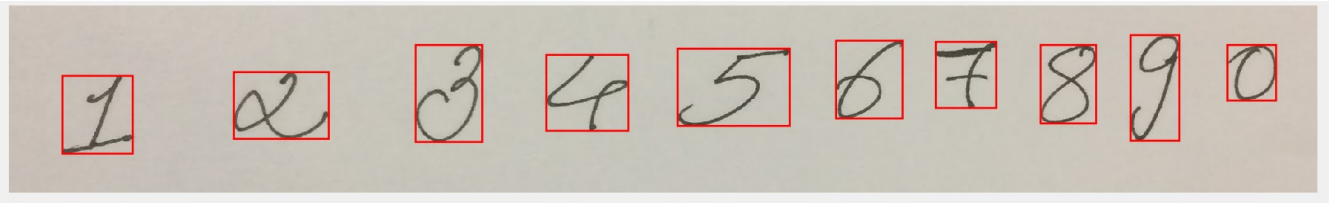


Image 2

```
prediction =

     5     8     5     3     5    10     4     5     8     6
```

For image 2, the binding box pick up the correct number of digits, but our network still only able to correctly guessed the number 5.

Image 3

```
prediction =

     9     5     4     4     3     8
```

Again there is an error on the predicted number of digits because the function find an extra connected components on the last number 6. This time the network was not able to guess any digits correctly.

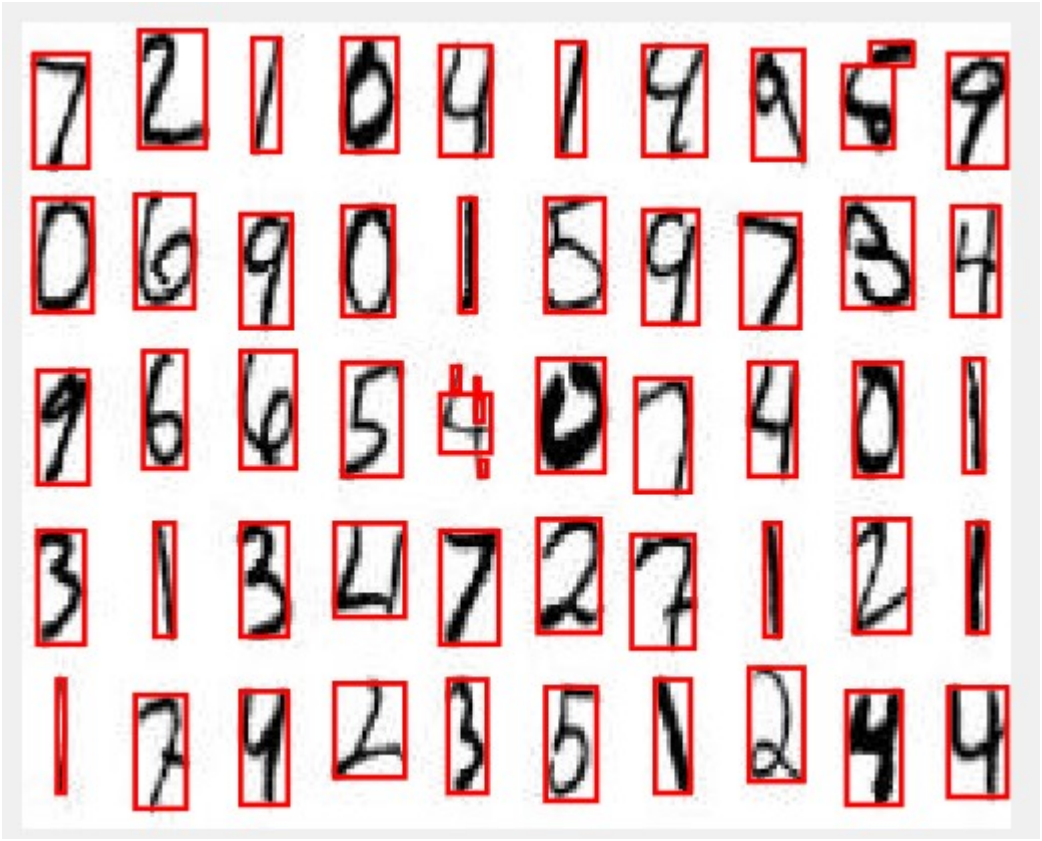


Image 4

```
>> ec

prediction =

  Columns 1 through 21

     8     9     8     8     3     9     8     8     4     3     6     9     4     6     3     3     8     4     8     4     8

  Columns 22 through 42

     3     6     8     3     3     8     2     8     4     8     4     8     3     8     8     8     8     3     6     4     6

  Columns 43 through 55

     4     3     4     8     8     8     8     6     5     3     3     8     8
```

There is an extra one bounding box on digits 8 (top right), and an extra 3 bounding box on digits 4 (in the middle), resulting an error in the number of predicted digits from our network.