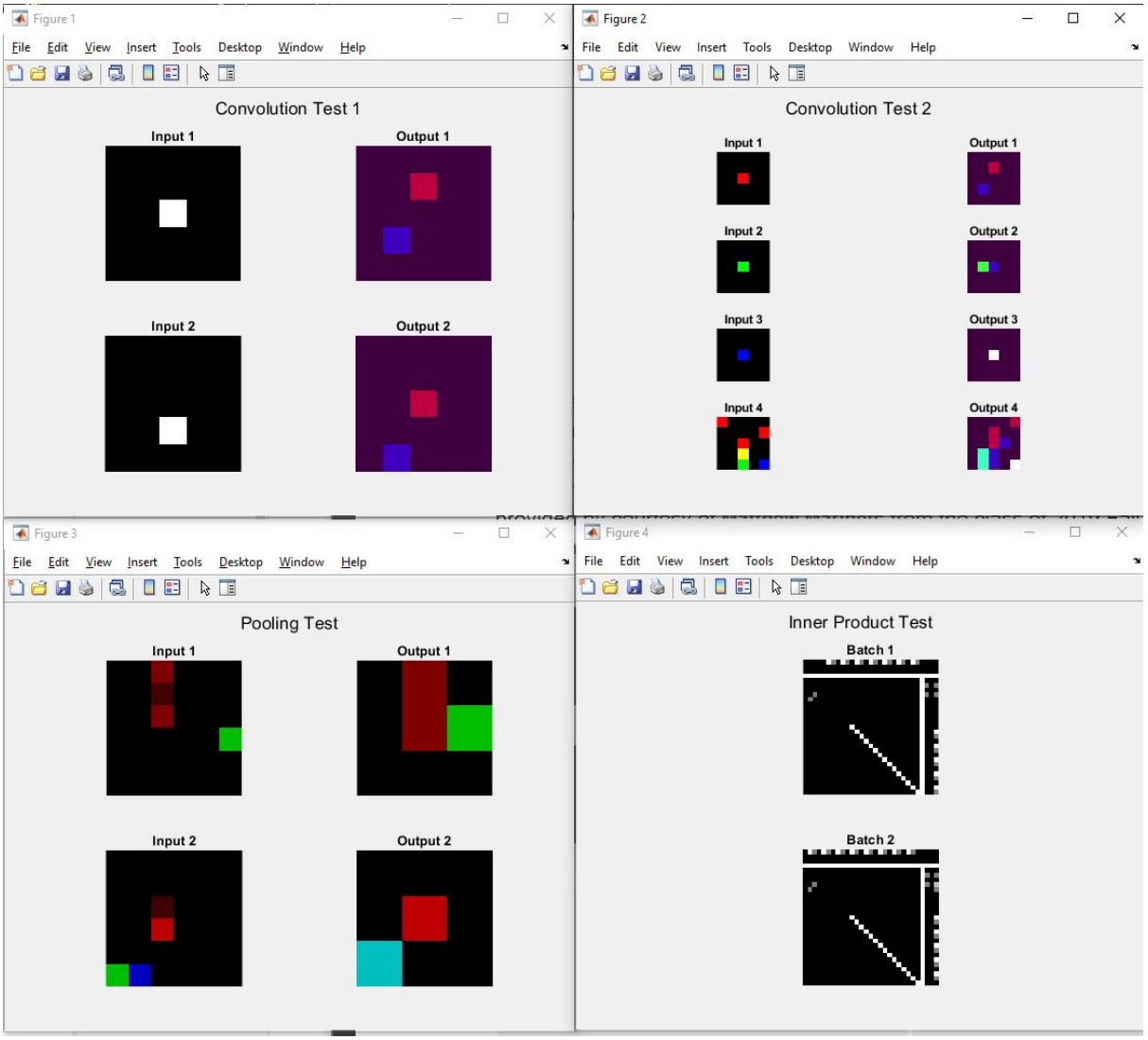
CMPT 762 Computer Vision

Digit recognition with CNN

For
Dr. Furukawa

By
BenKun Chen (bca96@sfu.ca)

1. Forward Pass



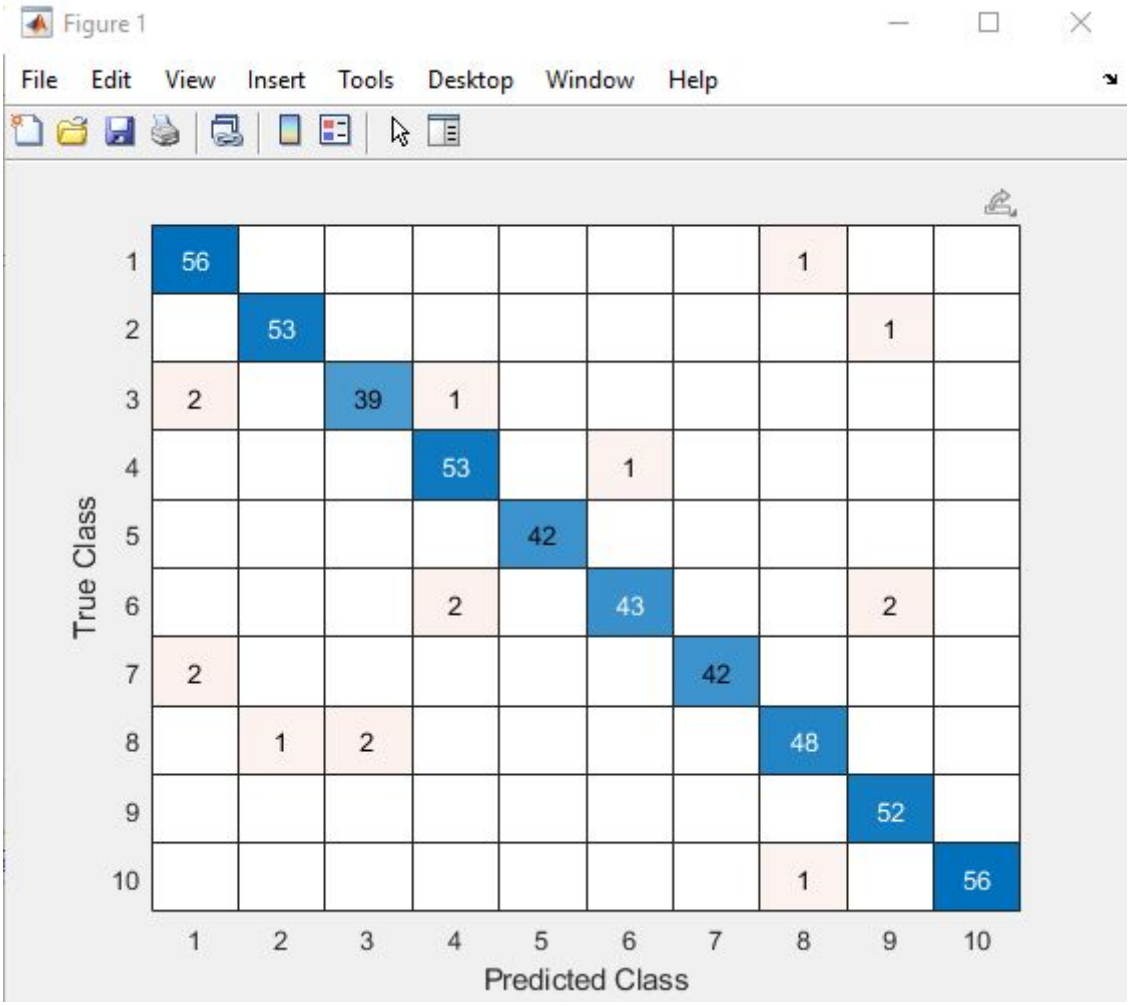
2. Training

a. Training

```
>> train_lenet
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

cost = 0.111115 training_percent = 0.960000
cost = 0.113216 training_percent = 0.940000
cost = 0.134874 training_percent = 0.960000
cost = 0.067548 training_percent = 0.990000
cost = 0.095426 training_percent = 0.980000
test accuracy: 0.966000

cost = 0.086685 training_percent = 0.980000
cost = 0.106186 training_percent = 0.950000
cost = 0.034245 training_percent = 1.000000
cost = 0.048397 training_percent = 1.000000
cost = 0.060728 training_percent = 0.970000
test accuracy: 0.968000

cost = 0.069977 training_percent = 1.000000
cost = 0.068312 training_percent = 0.980000
cost = 0.063643 training_percent = 0.980000
cost = 0.084625 training_percent = 0.960000
cost = 0.083214 training_percent = 0.980000
test accuracy: 0.970000

cost = 0.083081 training_percent = 0.970000
cost = 0.026531 training_percent = 1.000000
cost = 0.044653 training_percent = 0.980000
cost = 0.056298 training_percent = 0.980000
cost = 0.049833 training_percent = 0.990000
test accuracy: 0.970000

fx >>
```

The final accuracy is 95%

b. Test the network



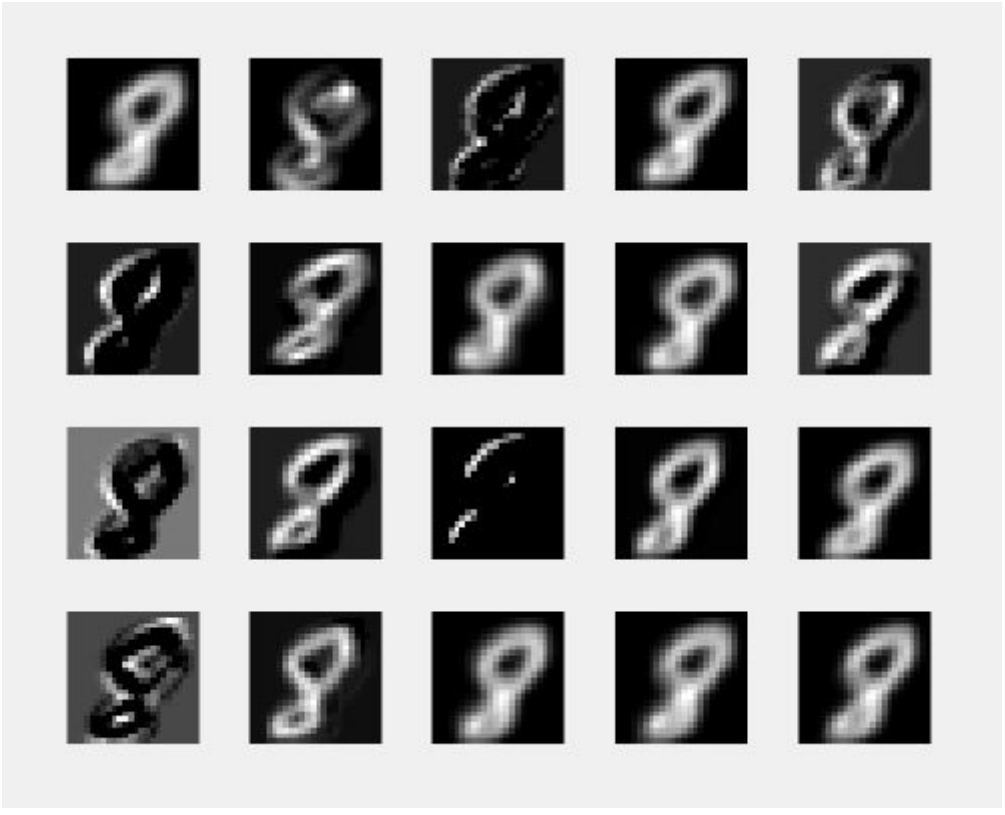
Predicted Class	True Class
3	8
1	7

They are confusing because 3 sometimes can look like 8 a lot depends on the curvy the handwritten digit is and 1 sometimes can be falsely recognized as 7 in some fancy handwriting styles.

c. Real -world testing


```
----------------
True Class:
     1

Predicted Class:
     1

----------------
True Class:
     5

Predicted Class:
     5

----------------
True Class:
     7

Predicted Class:
     7

----------------
True Class:
     8

Predicted Class:
     2

----------------
True Class:
    10

Predicted Class:
    10

----------------
>>
```

3. **Visualization**
    a. **CONV and ReLU layers**



    b. **Compare the feature maps to the original image and explain the differences.**

In the CONV layer each image has its concentration on features based on the bias & weight. We can observe that each image has its own highlights.
In the Relu layer the positive portion is updated more rapidly as training progresses. We can observe that it has increased the non-linearity in each image and results in a high contrast image.