

HW3 Report

Rui Huang, Zonglin Peng

The implementation is written in homework3_rhuang2_zpeng.py

Part One

Screenshot

The last 20 iterations of our gradient descent on the training data

```
In [595]: run homework3_rhuang2.py
Iteration: 80 | fCE: 0.19637158758274514 | fPC: 0.9496
Iteration: 81 | fCE: 0.19637002416148958 | fPC: 0.9496
Iteration: 82 | fCE: 0.1963686322051557 | fPC: 0.9496
Iteration: 83 | fCE: 0.1963673513060455 | fPC: 0.9496
Iteration: 84 | fCE: 0.19636622769521125 | fPC: 0.9496
Iteration: 85 | fCE: 0.19636515724752573 | fPC: 0.9496
Iteration: 86 | fCE: 0.19636423375569695 | fPC: 0.9496
Iteration: 87 | fCE: 0.196363381734605 | fPC: 0.9496
Iteration: 88 | fCE: 0.19636259709378923 | fPC: 0.9496
Iteration: 89 | fCE: 0.19636189704809087 | fPC: 0.9496
Iteration: 90 | fCE: 0.19636126145819244 | fPC: 0.9496
Iteration: 91 | fCE: 0.19636069516521887 | fPC: 0.9496
Iteration: 92 | fCE: 0.19636017764258082 | fPC: 0.9496
Iteration: 93 | fCE: 0.19635970927571272 | fPC: 0.9496
Iteration: 94 | fCE: 0.1963592913670253 | fPC: 0.9496
Iteration: 95 | fCE: 0.196358905883357 | fPC: 0.9496
Iteration: 96 | fCE: 0.19635855576126637 | fPC: 0.9496
Iteration: 97 | fCE: 0.1963582478298807 | fPC: 0.9496
Iteration: 98 | fCE: 0.196357963120429 | fPC: 0.9496
Iteration: 99 | fCE: 0.19635770925599455 | fPC: 0.9496
Training Accuracy: 0.9496
Testing Accuracy: 0.8818
Training fCE: 0.19635770925599455
Testing fCE: 0.4183988844635668
```

Explanation:

In the screenshot shown above, we could notice that the cross-entropy cost of the machine is decreasing as we apply SGD more, meaning the machine has better accuracy. But as shown above, the alteration of each step is really small, so consequently improvements in correctness is hard to detect in the last 20 iteration.

Here we achieved 94.96% accuracy on training set and 88.18% accuracy on testing set

Part Two

Sample Output

```
> run homework3_rhuang2_zpeng.py
Training Accuracy: 0.9506
Testing Accuracy: 0.8812
Training fCE: 0.1967206822472633
Testing fCE: 0.41906518816433835
Augmented Training Accuracy: 0.9095714285714286
Augmented Testing Accuracy: 0.8858
Augmented Training fCE: 0.32420493248511156
Augmented Testing fCE: 0.3912874735959556
```

Explanation:

In the data augmentation part, we used 7 augmented data sets:

- Translate up by 1 pixel
- Translate down by 1 pixel
- Scale the image up by 0.08 (*1.08)
- Scale the image down by 0.08 (*0.92)
- Rotate the image CCW by 5 degree
- Rotate the image CW by 5 degree
- Apply random noise on each pixel with normal distribution

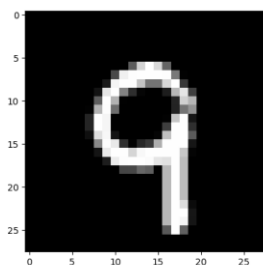
(Images are on the next page)

With the augmented training data set, the training accuracy decreases from 95.06% to 90.95% but the testing accuracy increases from 88.12% to 88.58%, which shows the machine has better performance using augmented data.

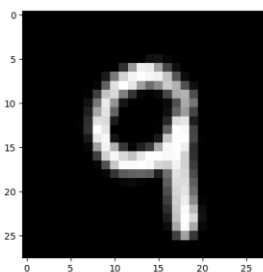
In line 140, change **augment = False to True to toggle the data augmentation*

Augmented data visualizations

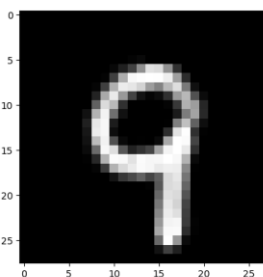
1. Original



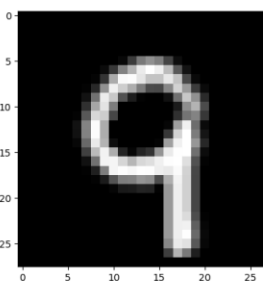
2. CCW 5 degree



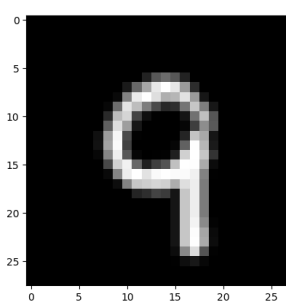
3. CW 5 degree



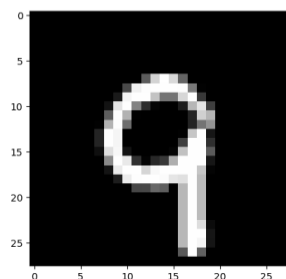
4. Scaled up



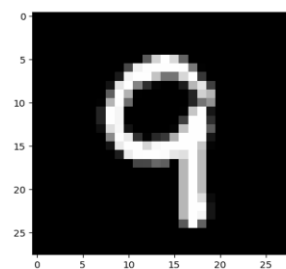
5. Scaled down



6. Translate down



7. Translate up



8. Noise:

