Homework 3

Yasheng Sun

[The project code is on my github https://github.com/sunyasheng/Neural-Network-Assignment]

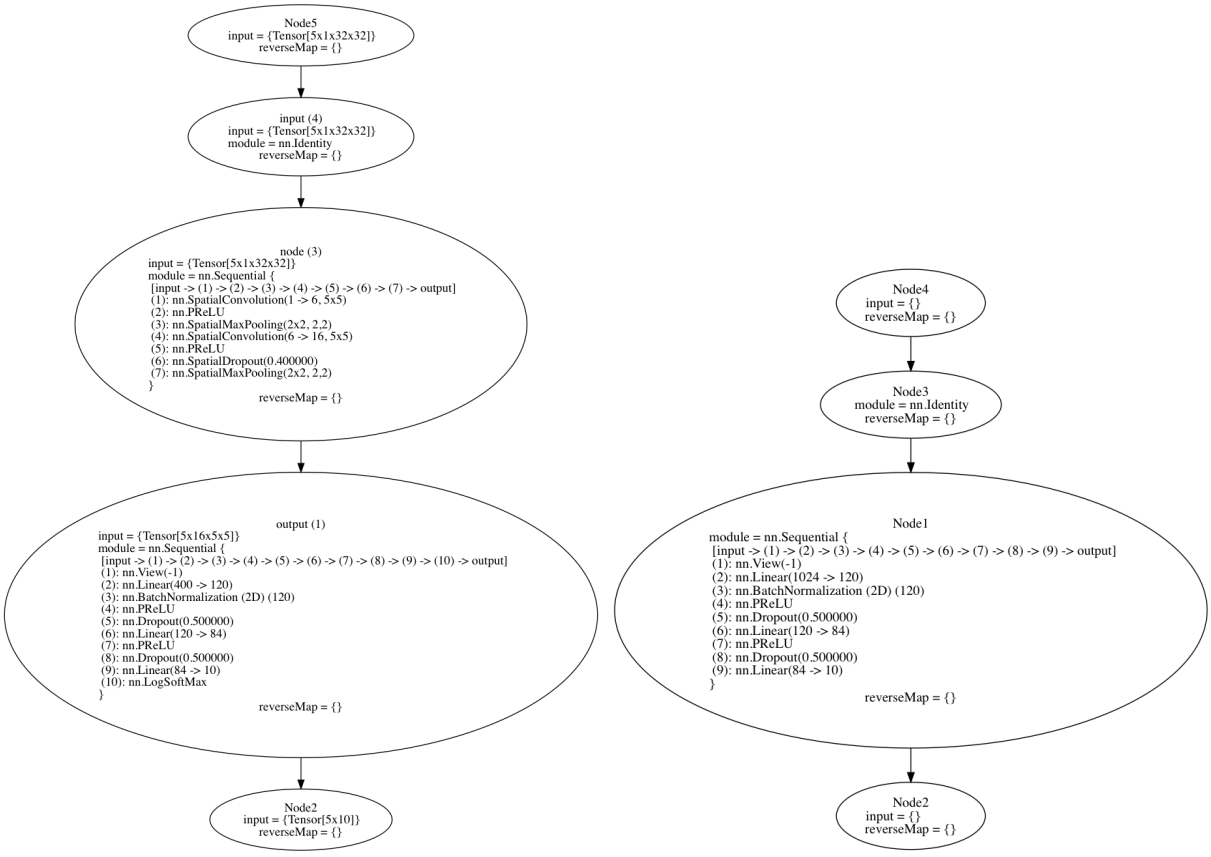
1 Solution

In this assignment, we compared the capabilities of feedforward neural network and CNN.

1.1 Network Structure

The feedforward neural network and CNN structure are shown below. Left is for feedforward network, right is CNN.

Graph model for feedforward network and CNN model.



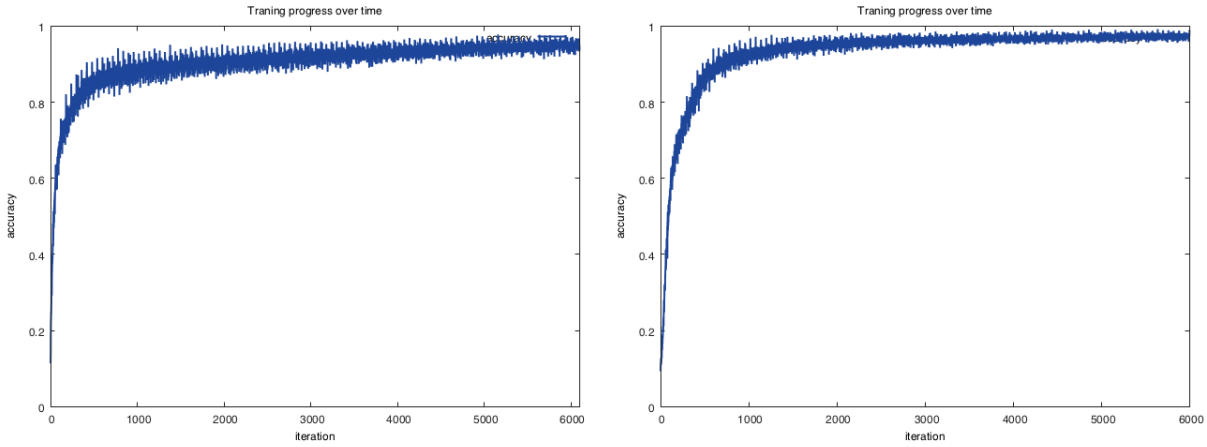
1.2 Training Time

The training process runs on Mac OS 10.12.6 with 2.5 GHz Intel Core i5 CPU and 10 GB 1600 MHz DDR3. For the feedforward structure, it takes 6 minutes to run 6000 iterations while it take 40 minutes for CNN.

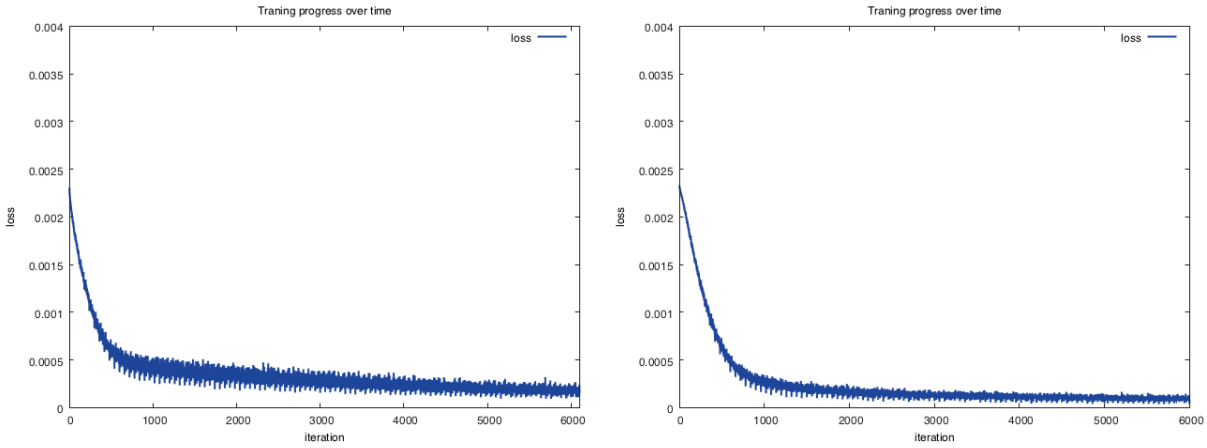
1.3 Training Error and Accurary

Both of these two networks perform well in test dataset. It seems that the CNN behaves more stable in training process. This is because CNN have the sharing weights technique which avoids overfiting to some degree. Left is for feedforward network, right is CNN.

Evolution of accuracy on training dataset in 6000 iterations.



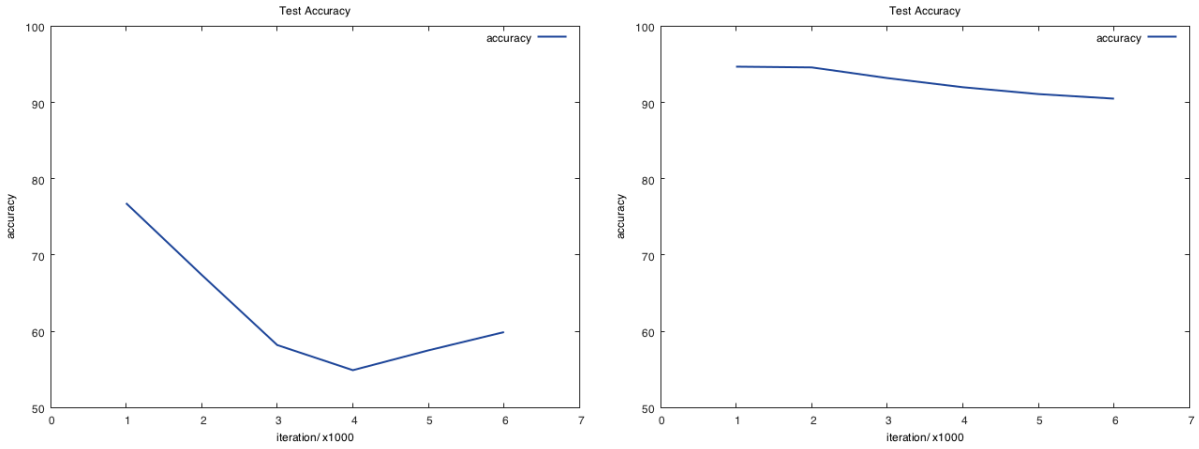
Evolution of CrossEntropy loss on training dataset in 6000 iterations.



1.4 Testing Error

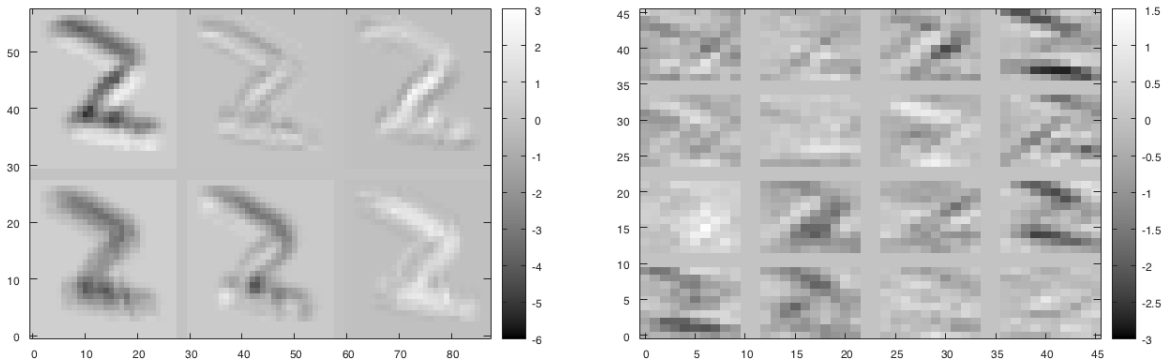
We test the trained model in testing dataset at 6 stages. The test accuracy is shown below. Left is for feedforward network, right is CNN. Obviously, the CNN performs much better than feedforward network. For CNN, the accuracy decreases with iteration, which may be caused by overfiting in training stage. For feedforward network, there must be some magic behind there. God knows why the accuracy decreases first and then increases gradually.

Evolution of accuracy on testing dataset at 6 stages.



2 Feature Visualization

Extracted Features are shown below. Left is the feature extracted in lower layers while right is the feature extracted in higher layers. It seems that the extracted features keep the whole shape of the original image.



3 Note

The usage of the code is illustrated in ReadMe which is also on my github.