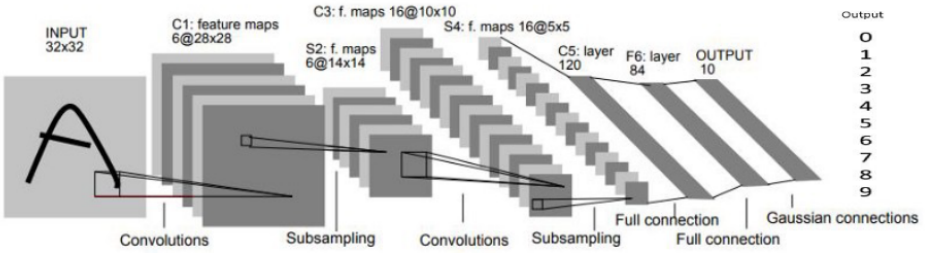
Assignment #2: Image Classification with LeNet-5

Hanyang University

Sooyoung Kim[redacted], Gaon Choi[redacted]

Introduction

LeNet is a neural network architecture for hand-written and machine-printed character recognition. In this experiment, the main goal is to train the LeNet model on MNIST dataset and to improve its accuracy upon 95%. The detailed architecture of LeNet is as follows.



i) First Layer: convolutional layer with 6 feature maps (filter size: 5 x 5)

ii) Second Layer: sub-sampling layer (average pooling, max pooling, etc.)

iii) Third Layer: convolutional layer with 16 feature maps (filter size: 5 x 5)

iv) Fourth Layer: sub-sampling layer (average pooling, max pooling, etc.)

v) Fifth Layer: Fully Connected Layer (120 84)

vi) Sixth Layer: Fully Connected Layer (84 10)

vii) Output Layer: 10 possible values corresponding to the digits from 0 to 9

Experiment

1. Model Architecture

We decided to use the proposed Model Architecture as it is.

The MNIST dataset is 28*28 size, which is not very large, and it has only one channel.

So, it was thought that two convolution layers were sufficient to extract features.

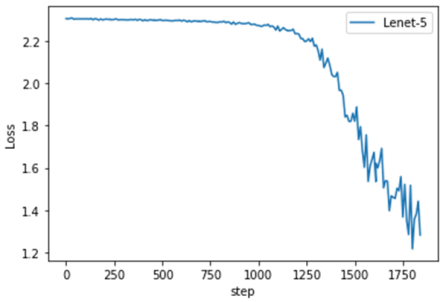
2. Training and performance

PyTorch's **CrossEntropyLoss** was used as a cost function used during learning, and PyTorch's **SGD** was used for optimization.

The CrossEntropyLoss function has parameters such as weight, size_average, ignore_index, reduce, reduction, and label_smoothing, but in this example, only default values are used.

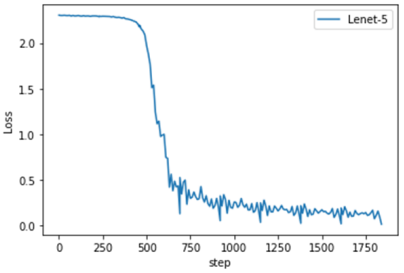
The SGD function has parameters such as learning rate, momentum, weight_decay, dampening, nesterov, and maximum, and in this example, learning rate and momentum are used.

3. Analysis of Hyper-parameter

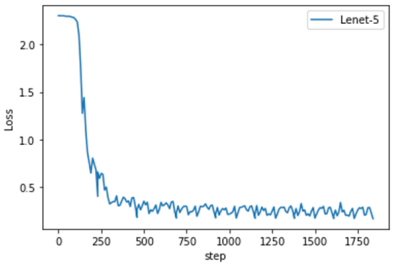


The results before parameter optimization are as follows.
The final loss was 1.28302 and the accuracy was 0.54. This was illustrated as follows.

As the first attempt at optimization through parameters, an attempt was made to change the learning rate. The learning rate was increased from 0.01 to 0.03, and the final loss was 0.01216, and the accuracy was 0.95. This was illustrated as follows.



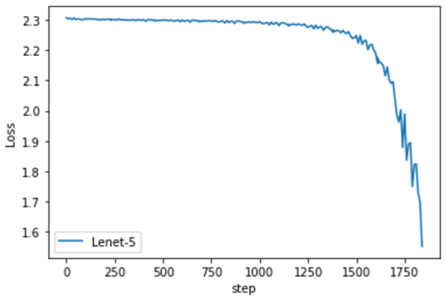
As you can see in the picture, oscillation occurs from 1,000 steps or higher. Therefore, rather than lowering the learning rate. It was intended to replace the epoch by increasing it.



Next, the parameter of SGD, momentum, was changed. After referring to several documents, 0.9 was judged to be appropriate, and The final loss was 0.28957 and the accuracy was 0.89. This was illustrated as follows.

The loss was seen to decrease rapidly, but couldn't reach the target value of 95%.

Finally, the experiment was resumed after increasing epoch from 8 to 20, and It exceeded the target of 95% accuracy and showed 99% accuracy. This was illustrated as follows.



However, it showed a difference whenever run it again, so could be expected to depend on the initial parameter setting.

Conclusion

Through this experiment, we implemented a deep learning model of LeNet structure for MNIST Dataset, a 28*28 size grayscale image, and also explored how to optimize it through several methods.

In conclusion, the best results were obtained at **learning rate 0.02, momentum 0.9, and epoch 20.**