

Financial Technology: Assignment 3

Dec. 2020

B07703014 財金三 蔡承翰

[Problem 1]

We have designed a CNN model using *PyTorch* package for this task. The model consists of three convolution layers and a fully connected layer. Each of the convolution layer is followed by a ReLu activation function, and a adaptive average pooling is placed behind the final convolution layer. Each of the convolution layer is of 16 filters. The visualization of our model is as follows.

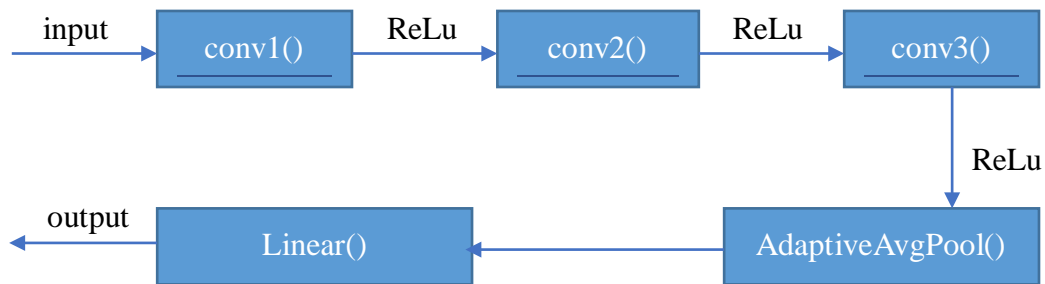


Figure 1: CNN model

The number of iterations is set to 25 manually. We have complete grid search on the stride size and the filter size. We've tried stride size of 1 and 2, with filter size of 1, 3, and 5. It turns out that the larger the stride size and the filter size are, the model gives higher accuracy on the validation set. Hence our best performing model is with stride size of 2 and filter size of 5. The model provides accuracy of 0.911 on the train set and 0.89 on the validation set.

[Problem 2]

Training the best performing CNN model, we are able to plot the following curves.

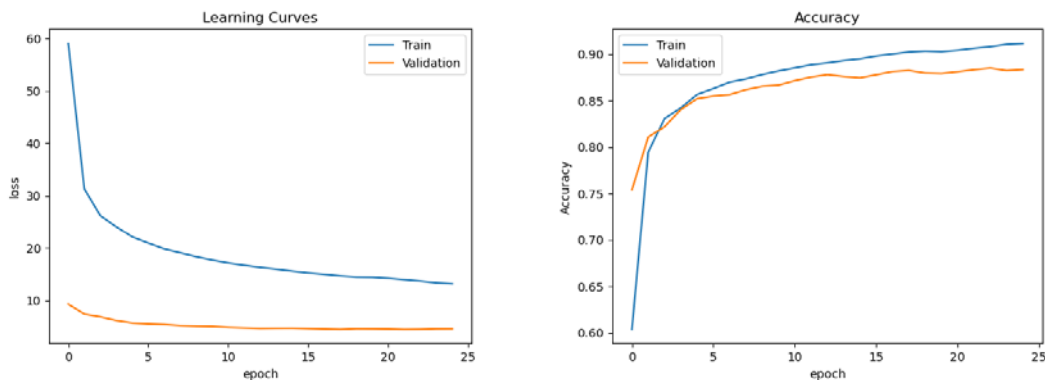


Figure 2: Learning curves and accuracy curve of CNN

[Problem 3]

The first image from the training set is shown as figure 3.

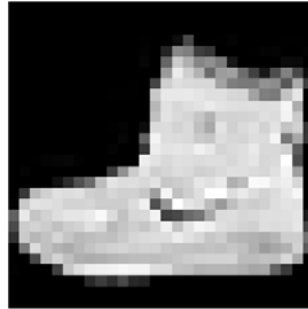


Figure 3: First image in the training set

The true classification of this image is category 9, which stands for ankle boot. After training the model, we have successfully predicted the true Y. The activations of the first convolution layer are shown in figure 4.

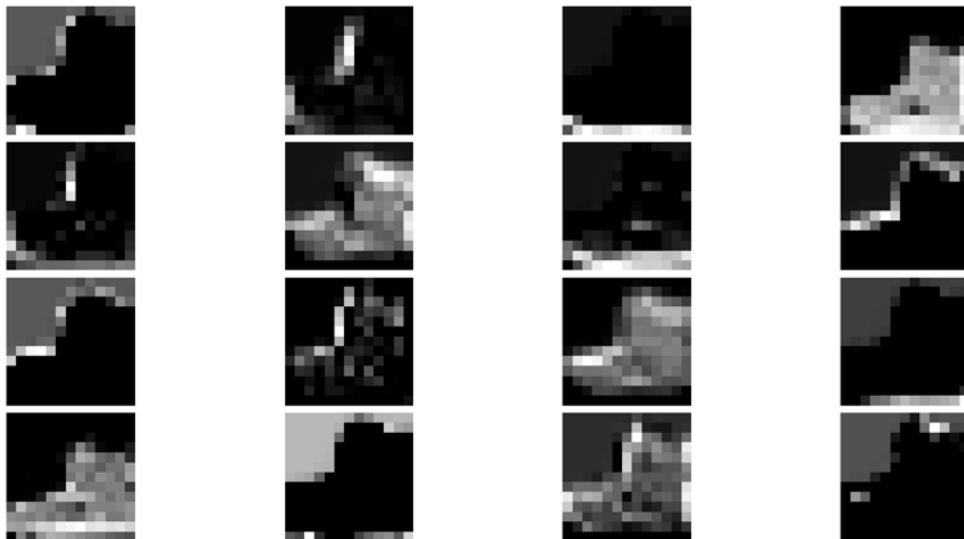


Figure 4: activations of the first convolution layer of CNN

We can see that each filter captured different parts of the images. Together with the following layers, the model is capable of classify the cloths using these features.

[Problem 4]

Despite we can now plot the prediction for every image in the validation set, notice that it is impossible for us to put it on this report as there are $0.2(60000 + 10000) = 14000$ images in the validation set. Instead, we plot the prediction and the images of the first 16 images in the validation set, which is shown in figure 5.

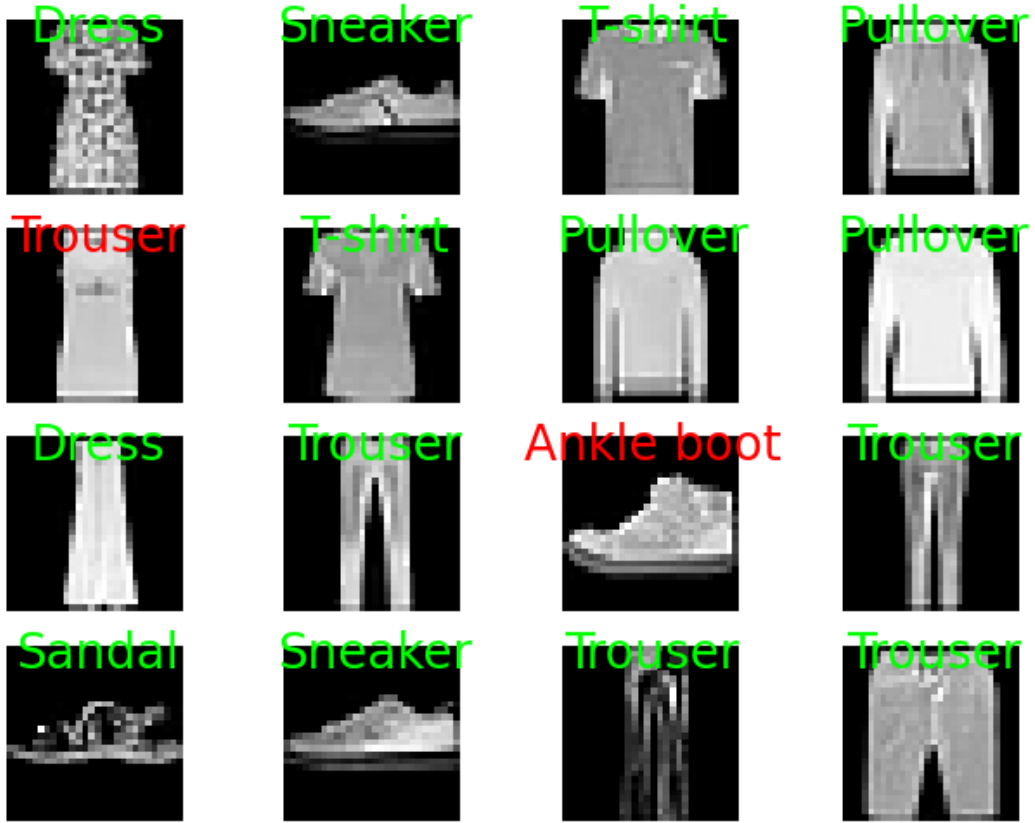


Figure 5: Predictions of CNN

[Problem 5]

We choose resnet as the improved model of CNN. Here we introduce *torchvision* package and train a resnet model with 18 layers and 64 filters. The model presented train accuracy of 0.967 and validation accuracy of 0.903, performing better than the regular CNN model we have introduced before. Number of iterations remains 25. The learning curves and the accuracy curves are shown in figure 6.

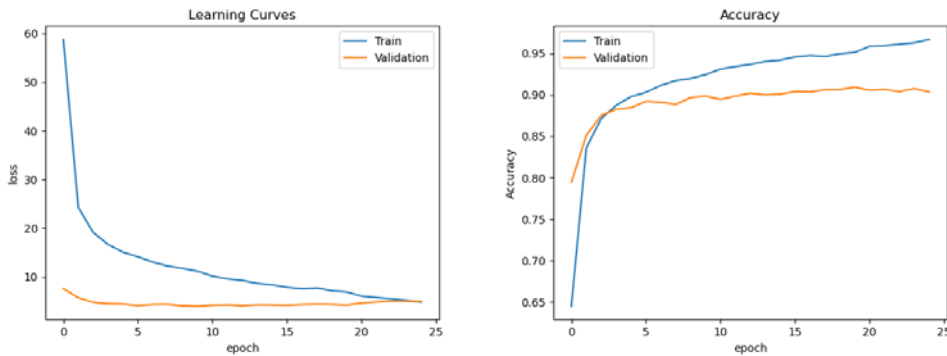


Figure 6: Learning curves and accuracy curves of resnet

The first image of the training set has been shown on figure 3. We present the activations directly in figure 7. The model classifies the image successfully.

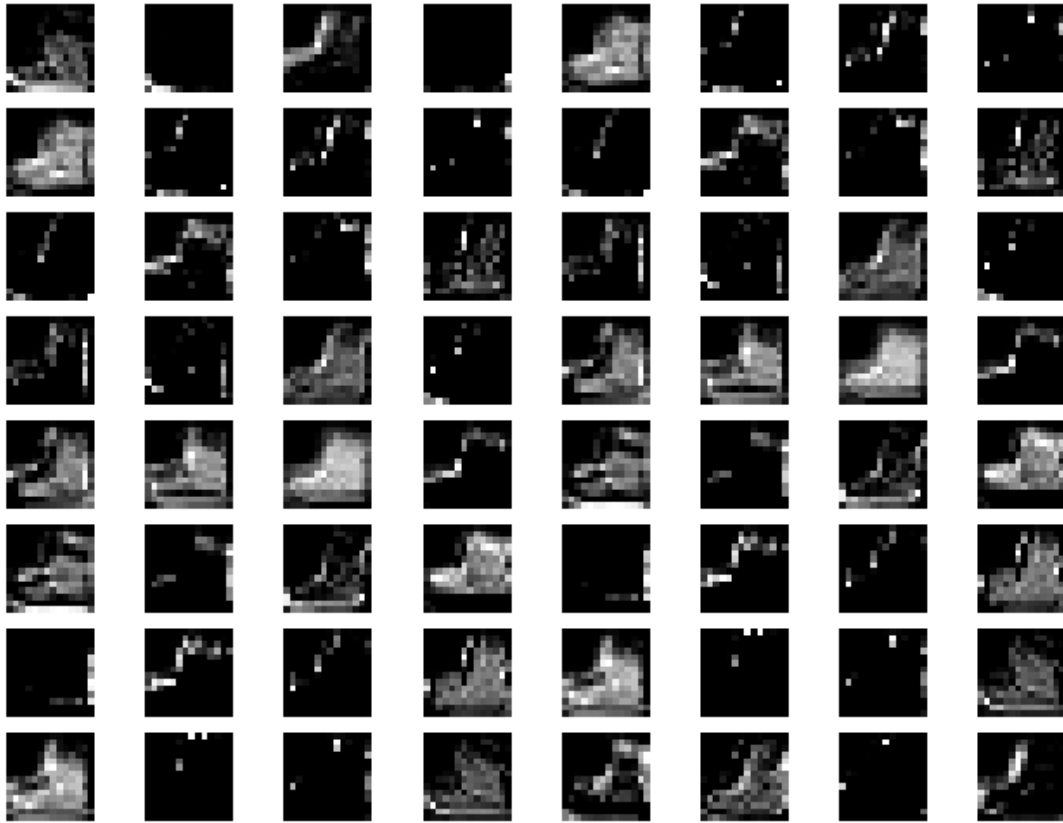


Figure 7: activations of the first convolution layer of resnet

The improved model has more filters, implying that it can capture more features of the image. Notice that not necessarily all the filters are meaningful. Some filters that cannot capture any feature in this image may be very helpful to classify another image. Finally, our classifications for the first 16 images from the validation set are shown in figure 8.

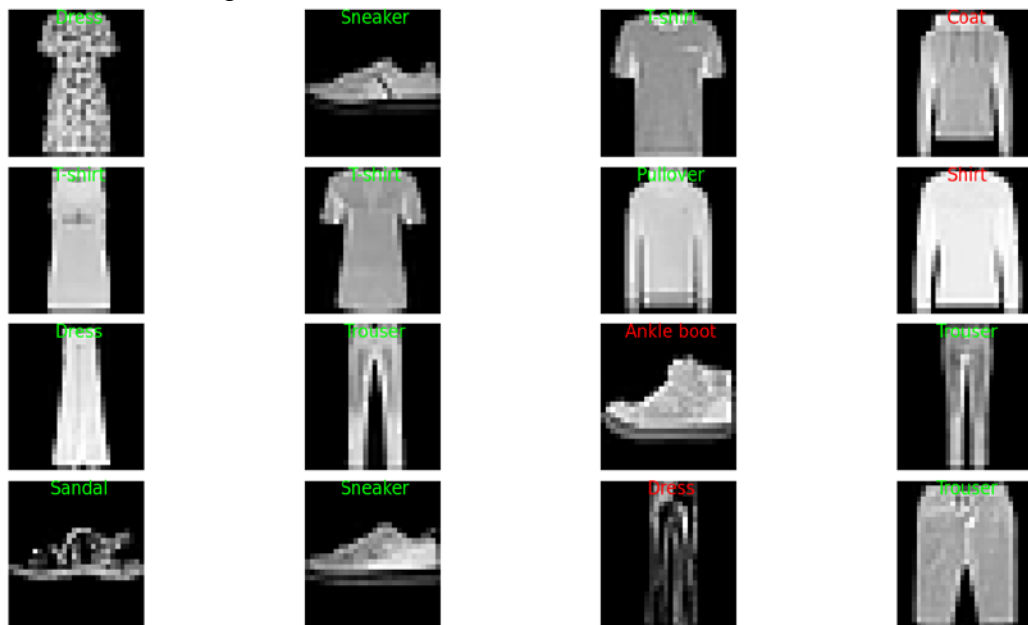


Figure 8: predictions of resnet

Though comparing figure 5 and figure 8, one may conclude that CNN model is better than resnet model, we shall focus on the total accuracy on the validation set. It is clear that the resnet model performing better on this task. Furthermore, comparing figure 2 and 6, we can see the difference between the training processes of the two models. In a nutshell, we must conclude that resnet performs better than CNN.