AI Couserwork Report

1. Experiment

(1) The model is built on required architecture of convolutional neutral network and its parameters.

a) CNN with 2 convolutional layers

After reducing the number of convolutional layers to two, the accuracy of the model has a slight growth, which is 49%. And the confusion matrix is produced as below:

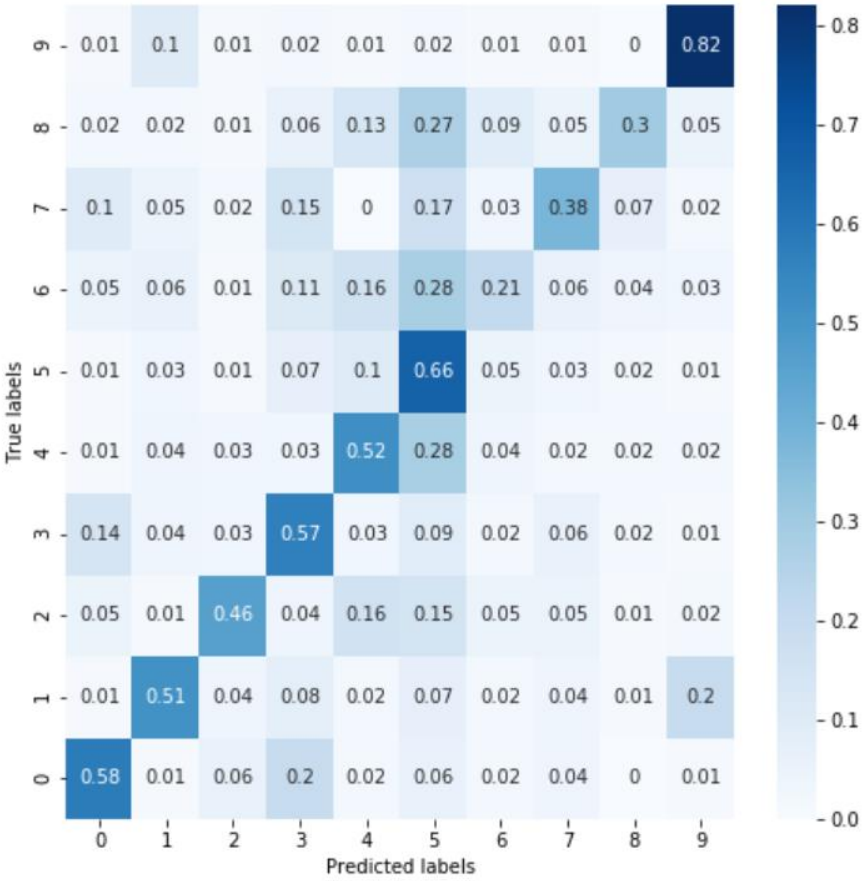

Figure 1. 2 convolutional layers

In this model, class 9 has the highest accuracy(82%), while class 6 is the lowest(21%), and most of the results on each class are around 50%.

b) CNN with 3 convolutional layers

This is the model totally based on the required architecture in the PDF file, and it finally turn out 43% of accuracy. And the confusion matrix stating each classes is below:

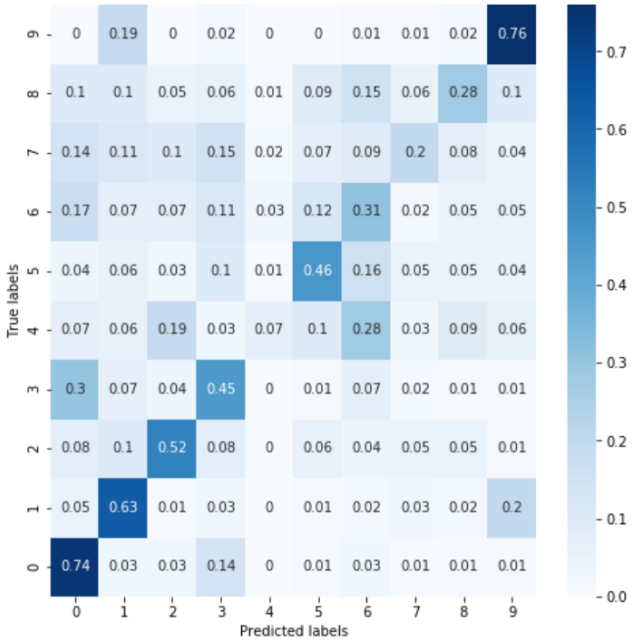

Figure 2. 3 convolutional layers

From the confusion matrix, we can see that the accuracy of class 0 and class 9 are the highest, both beyond 70%, and the accuracy of class 4 is the lowest, which has only 0.07. Therefore, class 4 is the main factor that make the entire model low accuracy.

c) CNN with 4 convolutional layers

When the dataset train in a model with 4 layers of convolutional neutral network, the accuracy on test set didn't increase, but decrease a little, it was 39%. And the confusion matrix shows below:

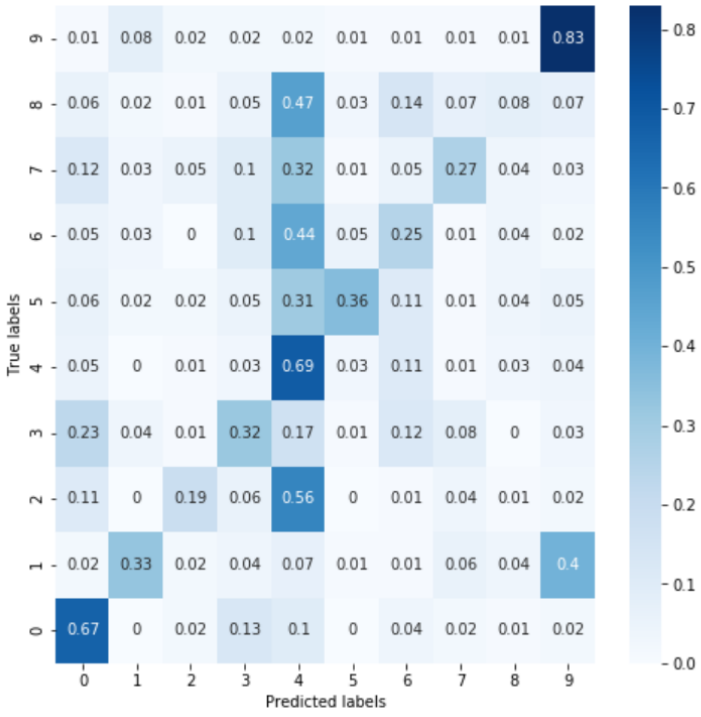

Figure 3. 4 convolutional layers

From the graph, we can see that class 9 still has the highest accuracy , only the accuracy of class 0, class 4 and class 9 are higher than the accuracy of the entire model. Therefore, more number of convolutional layer do not necessarily represent having a higher accuracy.

d) CNN with 5 convolutional layers

Having the largest number of convolutional layers, but the accuracy of this time is the lowest, which has only 28%.

Accuracy of the network on the test images: 28 %

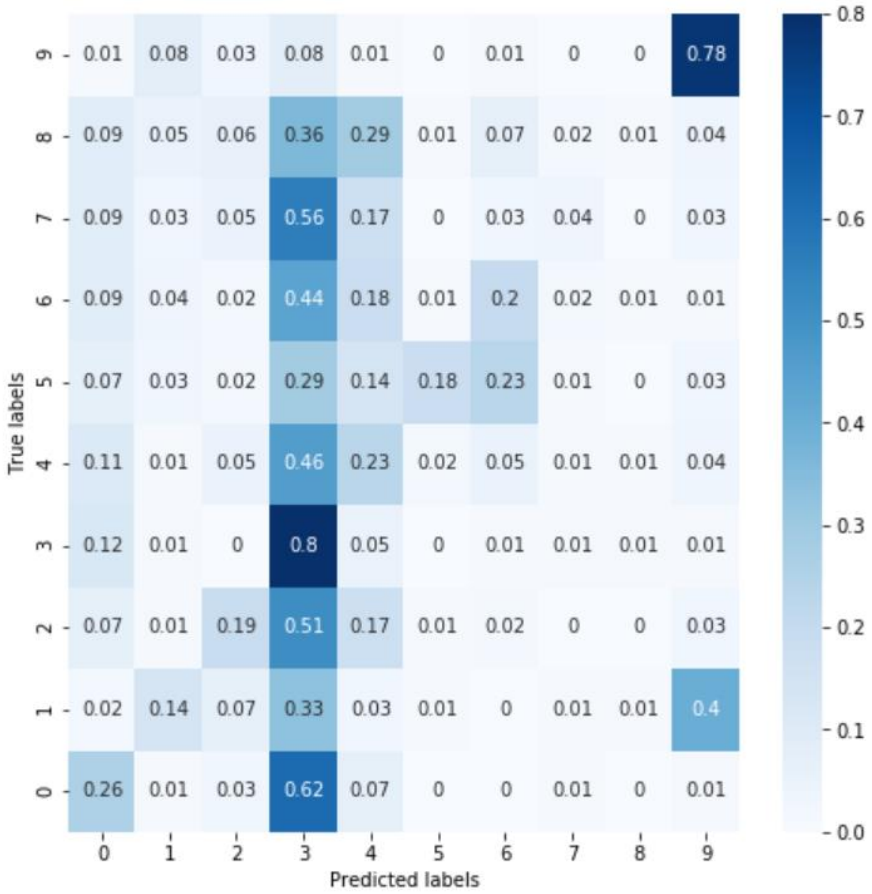

Figure 4.5 convolutional layers

2） Experiment with different architecture element

From the neutral network with different number of convolutional layers tested above, take the best performing one, which has 2 convolutional layers. Then I try to change the size of kernel and see what happen to the accuracy.

First, I changed the filter size of the second convolutional layer to 5*5, which is larger than previous size. After training process, we can see that the accuracy on test set decreased to 46%. From the confusion matrix graph shown below, we can see except there is class 9 always has the highest accuracy with 83%, and followed by class 0 with the accuracy 58%, other classes' performance are relatively bad, none of them exceed 46%. But at the same time, it doesn't exist that the accuracy of one particular class is super low. Therefore, changing the filter size to be larger may lead to a relatively average model, with no extreme results.

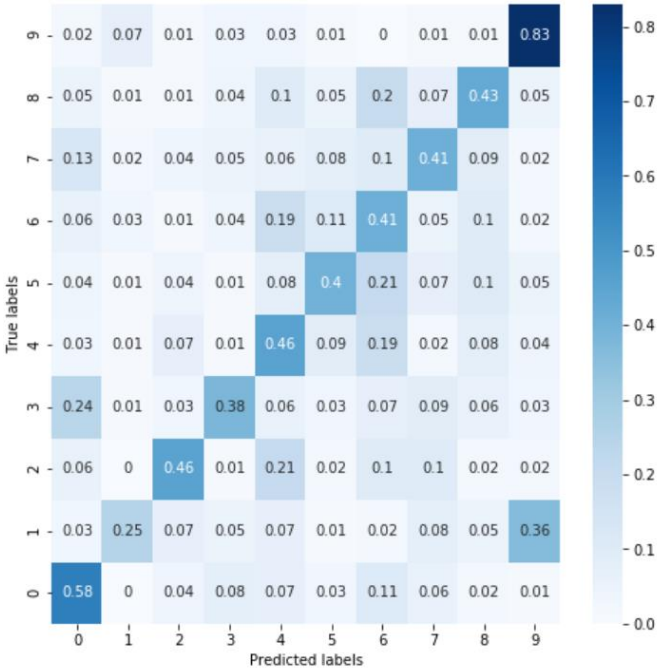

Figure 5. 5*5 kernel size

Subsequently, I changed the filter size to be lower than before. I changed it to be 3*3, and then the accuracy has 47%. Although still not better than 4*4, it is better than 5*5. From the matrix we can see more than a half of classes are more accurate than 47%. But because there are some classes with extreme low accuracy, the whole accuracy of the model is not so high.

And it is surprised that many class 9 pictures are classified to be class 1, which lead to such a low accuracy on class 9.

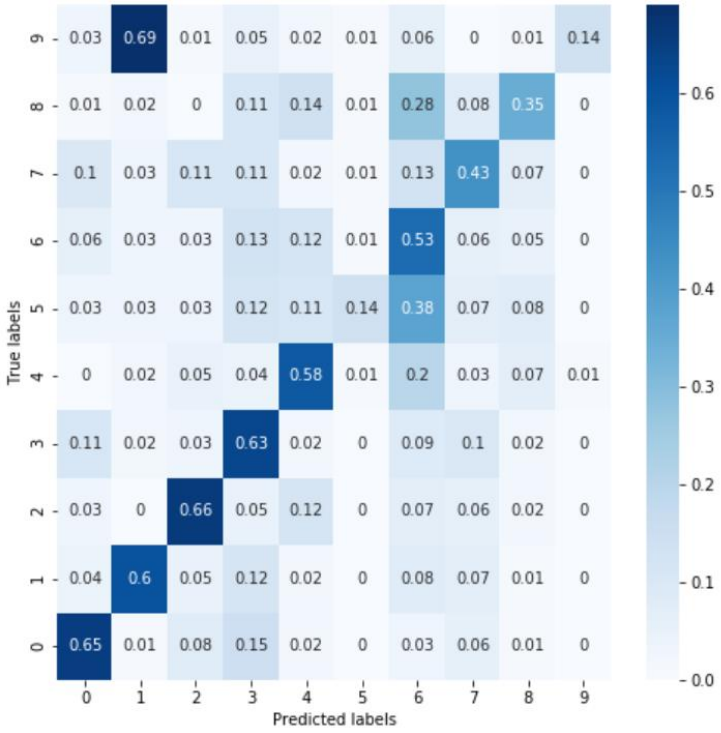

Figure 6. 3*3 kernel size

In summary, a larger size of filter will make the model more average, while small size of filter will make the model result more extreme.

2. filter visualization

Filters before training:

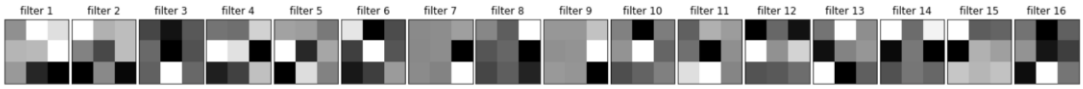

Figure 7. filters before training

Filters halfway during the training:

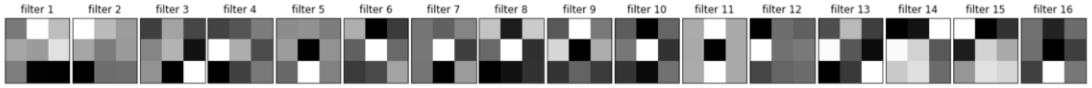

Figure 8. filters during training

Filters after the training:

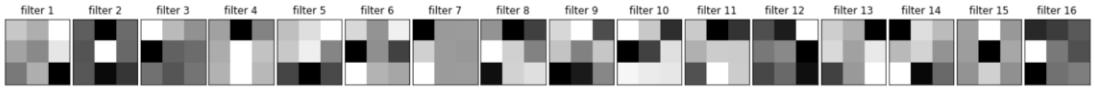

Figure 9. filters after training

For randomly generated convolutional kernel parameters, there is a large difference between the filters before training and the filters after training. The filters in training are similar to the filters after training in many places, and the color depth is a little different.

3. feature map visualization

In this section, we choose a group of pictures to test in the neutral network.

Figure 10. original test pictures

In the first layer, which has 3 channels for inputs and 16 channels for outputs. The pictures become slightly blurred in the first layer of neural network. And the classification of these pictures is [7, 5, 2, 6].

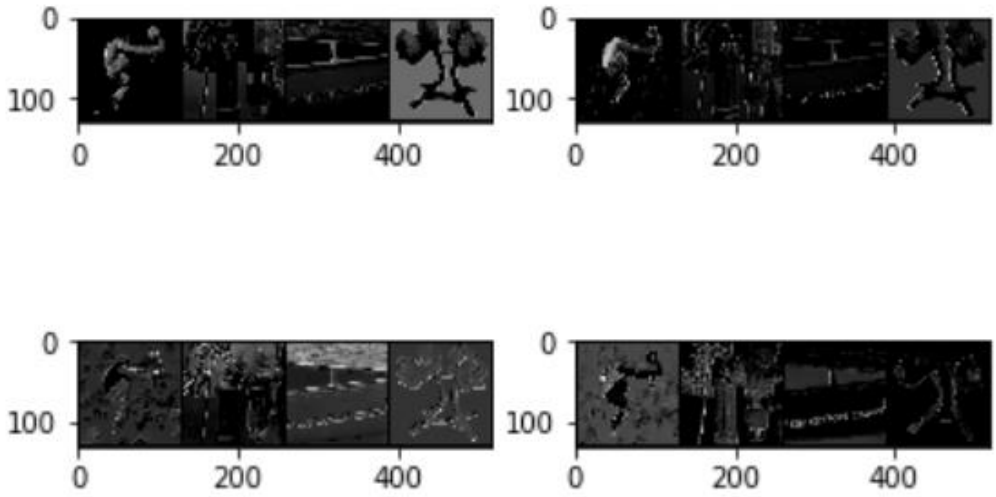

Figure 11. pictures after first layer

Form the pictures above, we can see that different classes of pictures have different feature maps after filtering. As we all known, the feature maps are generated from same number of channels, A channel is the detection of a certain feature, and the strength of a certain value in the channel is a response to the strength of the current feature. Therefore, the feature maps of the two pictures in different categories, the distribution of color depth is also different.

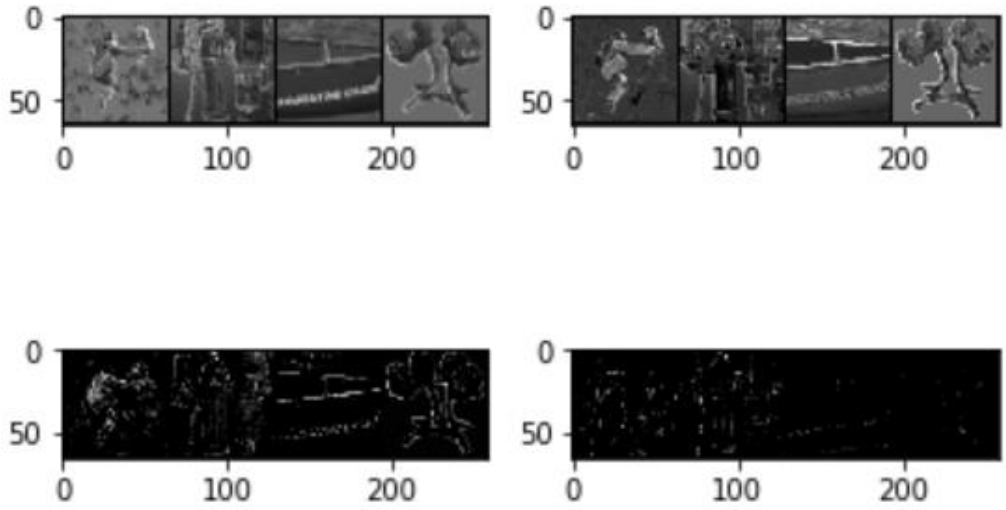

Figure 12. pictures after second layer

In the second layer of convolutional neutral network, pictures that in the feature maps getting much blurry. However, some of them become much lighter or darker.

4. Adjustment description

1） reducing the number of convolutional layers from 3 to 2.

After reducing one convolutional layer, the accuracy of the model on the test set and validation set has been improved. It can be known from the previous experiments that the accuracy rate of the neutral network with two convolutional layers can reach 49%, which

is the highest accuracy rate of all tests.

From my points of view, although deep learning neutral network requires more layers of neutral network, if you train too much on a certain data set, it will easily cause overfitting. By reducing the convolutional layers, you can prevent overfitting and improve the ability of generalization of the model.

However, getting such a result may just happen in my own experiments. If in another scenario with different images set and different neutral network architecture, then there will be a different story.

2) reducing the number of training time.

Although the higher number of training time can better adjust the weights of the neutral network and improve the classification accuracy of the entire training set. However, in this experiment, it does not mean that the more training time, the higher classification accuracy of the model. After different attempts, when epoch=2, the model will keep a relatively high accuracy and consume less training time. In addition, when training time of a model is very high, it is easy for a model to lose its generalization. Therefore, with a relatively low training time would make a model to be more generalized.