**Report of Experiment3&4**

1. **Date Processing**

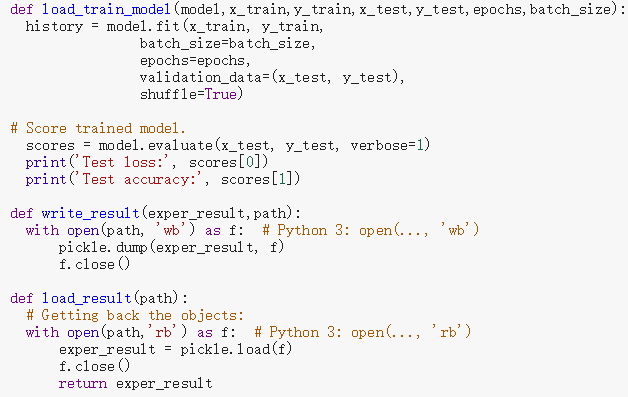
First, we use the dataset: tiny imagenet with 200 classes to train, validation and test. The shape of this dataset is 64, which is big so that would cost huge time, so we resized their shape as 32x32. We use the following code to reshape the dataset. 

Code Snippet 1

After resizing the data, we use following codes to load image data and separate them into training data, validation data. And I save loaded data as .npy file to save time. 

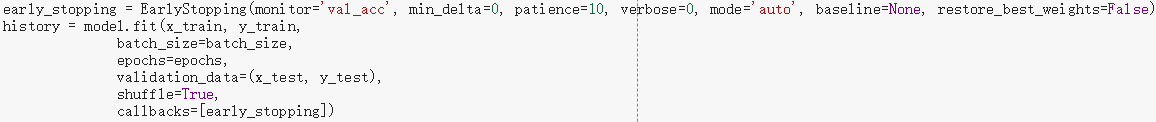
Code Snippet 2

1. **Utility and Early Stopping**

In order to save the training history and the result, I wrote a sort of functions to save my work. 

Code Snippet 3

Then I used early stopping to avoid my neural network from overfitting. In this function, I chose val\_acc to observe. If val\_acc is not improved, the model would not be trained. Here is the code of early stopping



Code Snippet4

1. **Experiment3**

I used Alexnet as my training model and did eight different experiments to observe how different combinations of parameters influence the result. The following code is the function of Alexnet model. 

Code Snippet5

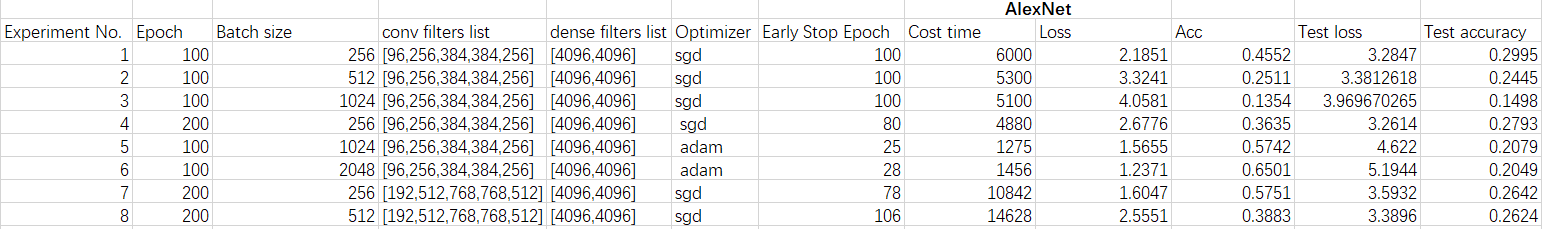
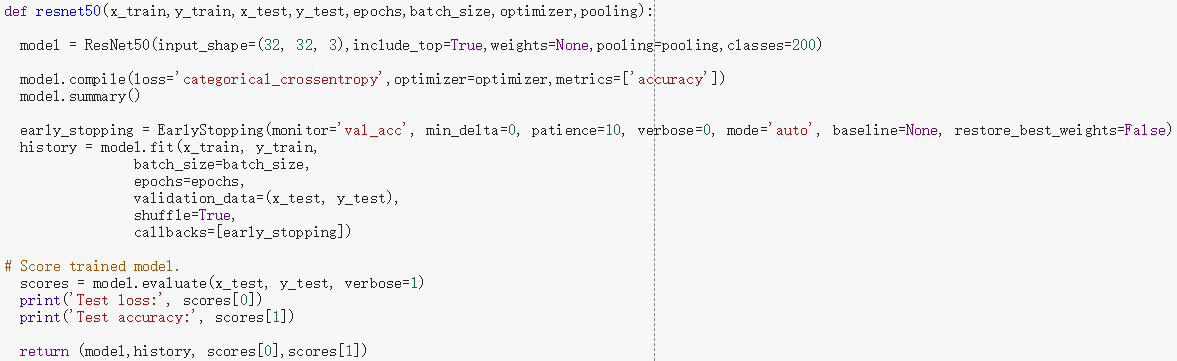
I used Excel to make the following comprehensive table which summarizes all experiments. 

Table 1

Through the table, we could find that the model of experiment 1 is the beset model. Although the training accuracy of this model is not the highest, the test accuracy is the best. Comparing Experiment1,2,3, we could find that the increase of batch size would result in worse test accuracy. The reason of this observation is that a larger would result in a significant degradation in the quality of the model and the generalization ability of the model [1]. Comparing Experiment3,5, we could find that adam optimizer could improve the accuracy of the model when the batch size is larger. In Experiment4,7, I doubled the number of filter, the model become deeper, but the result is not improved and cost more time.

1. **Experiment4**

I used Resnet50 as my training model and did six different experiments to observe how different combinations of parameters influence the result. The following code is the function of Resnet50 model.

Code Snippet 6

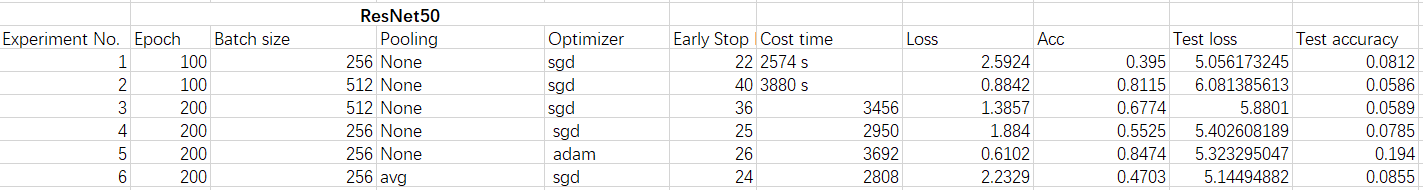
I used Excel to make the following comprehensive table which summarizes all experiments. 

Table 2

Through the table, we could find that the model of experiment 5 is the beset model. This model has the highest training accuracy and test accuracy. But the all of the model has the problem of overfitting. Comparing Experiment1,2,3,4, we could find that the increase of batch size would have the bad influence on the result. The reason is the same from above. Comparing Experiment 4,6, we could find that the test accuracy was improved after adding average pooling layer.

1. **Reference**

[1][Tradeoff batch size vs. number of iterations to train a neural network](https://stats.stackexchange.com/questions/164876/tradeoff-batch-size-vs-number-of-iterations-to-train-a-neural-network)