A0191508A Qiu Hai
CS5242 Neural Network and Deep Learning 2018/18 Sem 1
Assignment 2 Part 1

Task 1

We build the network according to the instruction, then start to compare the performance of different optimizer functions. All results are gathered with the same initial random weights, and trained for 20 epochs.

Optimizer	mnist		cifar	
	Loss	Accuracy	Loss	Accuracy
SGD	0.3116	0.9203	1.6351	0.4127
Adam	0.2675	0.9312	1.5713	0.4493
Adagrad	0.2218	0.9435	1.4738	0.4839
Adadelta	0.1920	0.9527	1.3854	0.5416

Table 1. Testing result after 20 epochs of given CNN

Here, we append the images of training result across 20 epochs for different optimizers. We can observe that SGD gave us the worst performance, and Adam, Adagrad and Adadelta almost gave us the same results. For a complex network, a constant learning rate is not a good strategy, an optimizer can adapt the learning rate according to the previous gradient changes will provide a better result.

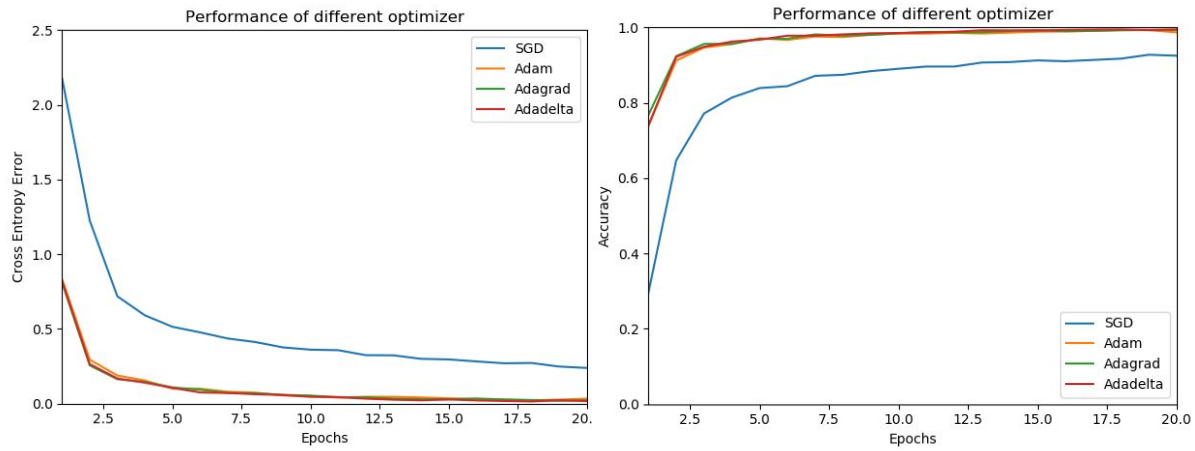

Figure 2. Training result of MNIST on different optimizer

Task 2

We build our own CNN, referring to the MINST cnn example provided by keras team [1]. The structure of the network is shown in the image below. I also append the prediction results of the trained network for both mnist and cifar10 test data, so we can visualize how this classification is being made. Compared to the CNN in task1, this network has less convolution and pooling layers, which maybe the cause of performance improvement, because the input size is 28*28

(32*32 for cifar), and each pooling layer will reduce the size by half, the final output size of the third pooling layer in task1 is 2*2, some features have already been lost in this process. For the dense layer after the flatten layer, we also choose a smaller output size (128), so the final sofmax output can be made base on a group of concentrated dominant features..

An interesting thing we observed is that in cifar, deer sometimes is recognized as bird or airplane, maybe the antler of the deer is extracted as the same feature as the wings of airplane or bird.

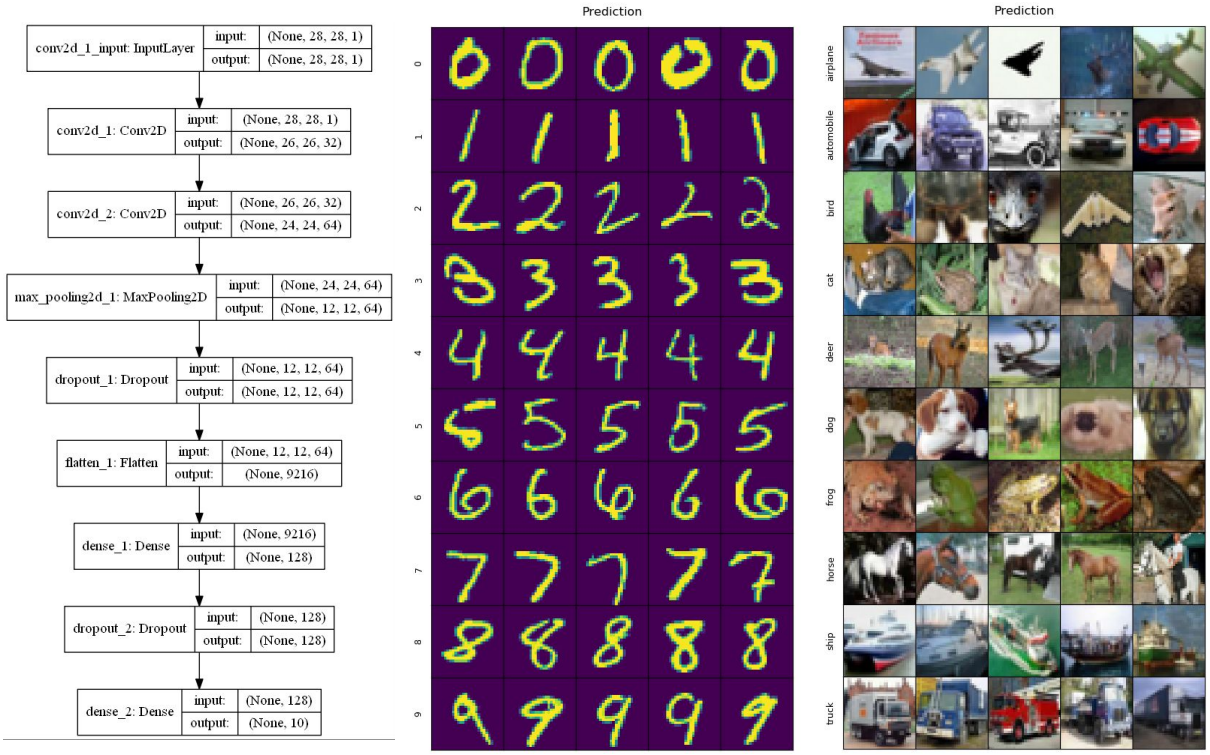

Figure 2. CNN inner structure and classification prediction result

Optimizer	mnist		cifar	
	Loss	Accuracy	Loss	Accuracy
SGD	0.1872	0.9544	1.1473	0.5928
Adam	0.0920	0.9762	0.8712	0.7918
Adagrad	0.0745	0.9845	0.4939	0.8592
Adadelta	0.0275	0.9907	0.1721	0.9495

Table 2. Testing result after 20 epochs of our own CNN

[1] https://github.com/keras-team/keras/blob/master/examples/mnist_cnn.py