Stat665 Lab7

Bo Song

I. Convolution model kernel size

Kernel size	1 * 1	3 * 3	5 * 5
Accuracy	0.91500	0.94700	0.93850

The 3 * 3 kernel performs best. The classification rate in using CNN performs better. However, I don't think there are any comparability between this pset and the previous one since there are 3 response classes in previous pset while there are only 2 in this one. And it's easier to get a higher classification rate in a small response classes situation. 3 * 3 kernel gives the best result. It may because this size can capture local features best. 1 * 1 is too small while 5 * 5 is too vague.

1 * 1 is an extreme case. We are essentially saying low level features are per-pixel, and they don't affect neighbouring pixels at all, and that we should apply the same operation to all pixels.

II. Convolution model as autoencoder

Kernel size	1 * 1	3 * 3	5 * 5
MSE	0. 02960	0.03169	0.03552

The MSE of 1 * 1 kernel is the smallest. It's because this kernel can preserve most information of original image. The dropout layer drops some information so the MSE is not 0.

III. Freezing the features

	Part1	Part2
Train Accuracy	0.98660	0.94110
Test Accuracy	0.94450	0.91500

The test accuracy is less than training accuracy, which is natural. The test accuracy of the model trained by autoencoder is not as good as that trained directly. We can guess that autoencoder is not suitable for CNN.

IV. Transfer learning

Model	Part 1	Part 2
Accuracy	0.66290	0.56300

The accuracy rate I achieved in problem set 5 is 51.72%. Though we only use 2 classes to train our CNN model, it still behaves better in 10 classes classification task than NN model in problem set 5.

V. Double convolution

The accuracy for double convolution model is 0.94700. Somehow it is equal to the 3 * 3 single convolutional model in part1. However, I believe this is just a coincident and double convolution should perform better than single one.

VI. Extensions

The model uses double convolution both with 3 * 3 kernel, followed by a 512 dense layer. The model is trained by 10 classes data set in CIFAR-10. The final accuracy is 0.71810. It cost about 4 hours to train the model. The model setting is the same as that in part 5 while we use full data set in training and get a better result.