

Theory and Practice of Deep Learning

Chang Jun Qing 1002088

Week 4

Task 1

Pretrain classifier used: Squeezenet v1.1

Based on prediction accuracy, we can see that with normalization the pretrain model works better and this is expected since normalization augments the existing data, make it more define for prediction. The following are the results.

Center Crop with no Normalize

Currently on Image 400

Currently on Image 800

Currently on Image 1200

Currently on Image 1600

Currently on Image 2000

Currently on Image 2400

Performance: Accuracy: 340/2500 (13.60%)

Figure 1: Center Crop With no Normalization

Center Crop with Normalize

Currently on Image 400

Currently on Image 800

Currently on Image 1200

Currently on Image 1600

Currently on Image 2000

Currently on Image 2400

Performance: Accuracy: 1442/2500 (57.68%)

Figure 2: Center Crop With Normalization

Task 2

Pretrain classifier used: Squeezenet v1.1

Comparing the results of five crop to just center crop, we see a slight increase in accuracy. Since the difference is not too big (about 4%), the difference could have either been due to the way the data was shuffled, or it could also have been that five crop is a better data augmentation than just center crop. The following is the results of five crop normalized.

Five Crop with Normalize

Currently on Image 400

Currently on Image 800

Currently on Image 1200

Currently on Image 1600

Currently on Image 2000

Currently on Image 2400

Performance: Accuracy: 1536/2500 (61.44%)

Figure 3: Five Crop Averaged With Normalization

When considering ten crop (five crop but with verticle mirroring), we can see how ten crop can be useful at times where certain images are clearer (easier to tell) when flipped. However, considering a case where an image is highly symmetrical (e.g. an image of an apple in the middle), it makes less sense since the images derived from the first five crop compared to the mirrored ones are very identical and would not value add.

Task 3

Pretrain classifier used: Squeezenet v1.1, InceptionV3

In both scenario, an adaptive average pooling layer was added to fit the input size of 330. We can see that inceptionV3 works better than squeezenet. The increase in number of parameters for inceptionv3 could have made it more accurate (however, the tradeoff is longer computation time)

Input Size 330 with SqueezeNet

Currently on Image 400

Currently on Image 800

Currently on Image 1200

Currently on Image 1600

Currently on Image 2000

Currently on Image 2400

Performance: Accuracy: 1509/2500 (60.36%)

Figure 4: Squeezenet with input size 330

Currently on Image 400

Currently on Image 800

Currently on Image 1200

Currently on Image 1600

Currently on Image 2000

Currently on Image 2400

Performance: Accuracy: 1863/2500 (74.52%)

Figure 5: InceptionV3 with input size 330