Deep Learning Assignment 1 Report

Kng Yew Chian

Introduction

A few experiments were conducted to observe the different approaches of transfer learning in a convolutional neural network. The datasets used were the Stanford Dogs dataset, and the Kaggle cats and dogs dataset.

In this submission, there are 5 notebooks:

1. experiment1a.ipynb : model trained on cats and dogs dataset, based on the tutorial mentioned in the assignment description.

2. experiment1b.ipynb : model trained on stanford dogs dataset. used as the base model for the next three experiments

3. experiment2.ipynb : stanford dogs base model, trained on cats and dogs dataset. Only the output layer was replaced.

4. experiment3.ipynb : stanford dogs base model, trained on cats and dogs dataset. Only the output layer and the first two convolutional layers were replaced.

5. experiment4.ipynb : stanford dogs base model, trained on cats and dogs dataset. Only the output layer and last two convolutional layers were replaced.

Background

Transfer learning refers to using pre-trained weights, usually from models trained on a similar dataset, to train a model on a new problem. In these experiments, we vary the amount of pre-trained weights used from the base model and observe the results.

The base model here refers to a convolutional neural network that has been trained on the Stanford Dogs dataset, which is used to classify between dog breeds. The new problem is to classify between cats and dogs, which is a binary classification problem instead, hence the need to replace the output layer of the base model.

Results

	Experiment 1a	Experiment 2	Experiment 3	Experiment 4
Validation accuracy (1 epoch)	0.4964	0.8813	0.5248	0.6149
Validation accuracy (25 epochs)	0.9423	0.9661	0.9545	0.9197
Epoch when 0.9 accuracy was	19	2	6	2

crossed				

Discussion

1. In general, the validation accuracy at 25 epochs for experiments 1a, 2 and 3 were quite similar. Experiment 4, which had its last two convolutional layers and its output layer replaced, seemed to have a lower validation accuracy than the other experiments.
    a. This difference might suggest that the weights in the last two convolutional layers take longer to train in the later than those in the first two convolutional layers. We may conclude that a model with this architecture is able to recognize and generalize lower level abstractions of an image such as edges at a faster rate in the later epochs than higher level abstractions such as more complicated shapes and combinations of features.

2. After the first epoch, experiment 2 had a much higher validation accuracy than the other experiments.
    a. Of all the experiments, experiment 2 had the greatest amount of pre-trained weights (all layers had pre-trained weights except for the output layer). This might suggest that the weights trained on the Stanford Dogs dataset are highly suitable for the problem of classifying cat and dog images.

3. After the first epoch, the validation accuracy of experiment 4 is slightly higher than experiment 3 (0.6149 > 0.5248).
    a. Since the weights have hardly been trained at the end of the first epoch, the difference between the two models in experiments 4 and 3 could mostly be attributed to the importance of the convolutional layers. Specifically, if the first two convolutional layers had bad weights, how much does that affect the models accuracy? The same question could be asked conversely. From the difference in the validation accuracy, it seems that the weights of the last two convolutional layers are not as important in providing a good accuracy.
    b. However, the weights of the last two convolutional layers might play a bigger role in pushing the validation accuracy past the 0.6149 threshold. Hence, we cannot generalize the importance of the last two convolutional layers with respect to training the whole model.

4. Experiment 4 reached a validation accuracy of 0.9 much faster than experiment 3.
    a. This might suggest that the last two convolutional layers train faster than the first two layers to reach a moderately high accuracy.

Conclusion

These experiments were conducted to evaluate the effectiveness of transfer learning. We assume that transfer learning is said to be effective when it allows a model to reach a good accuracy at a faster rate.

At low epochs, it seems that transfer learning is especially effective when replacing the last few convolution layers. We can guess that higher level features are quick to train at low epochs, possibly because the low level features such as simple shapes and edges are already recognized by the model, and it is simple to put together these features in a meaningful way. After about 6 epochs, the pre-trained weights do not seem to affect training rate or accuracy much.

At the very least, it seems like using the Stanford Dogs dataset to solve Kaggle cats and dogs dataset related problems seems to be useful and effective.