Deep Learning: Assignment 1 Convolutional Neural Network

Samuel Joseph Mcmurray

April 2023

The assignment is to conduct 4 experiments under the topic of computer vision, the task is a image classification using a Convolutional Neural Network(CNN) using the kaggle cats and dogs dataset as well as a the Stanford dogs dataset. The entire process follows the keras tutorial where the pre-processing is performed on the data by filtering for corruptions, generating the training and validation datasets, visualization, and augmentation to reduce overfitting. The CNN model architecture consists of an 32 hidden layers, the input layer, the rescaling layer, and the output layer. The hidden layers consist of 2-D convolution and separable convolution layers, as well as batch normalization, pooling, and activation layers.

1 Experiment One

The first experiment is consists of using the tutorial and the model to experiment on the Cats and dogs dataset with a learning rate of 0.0001, due to GPU limitations the batch size had to be reduced as the memory of the GPU was insufficient the original being 128 images reduced to just 32. The final epoch of the training set had a loss of 0.0902 an accuracy of 96.66%, while the validation loss was 0.2109 with an accuracy score of 93%. When compared to the original scores from tutorial there is an observed loss in performance of about 1-2% in the training in validation due to the change in the batch sizes being different.

The second part of the first experiment was to continue with the existing architecture to train a model on the Stanford Dogs dataset. The output layer was adjusted to increase the neurons to 120 classes corresponding to the number of classes in the dataset as well as a change in the scoring function which was changed to Sparse categorical Cross-Entropy to handle the additional classes. The epochs and learning rate were unchanged as no mention of changing these was mentioned in the assignment but given the results a longer epoch of maybe 100-150 would have seen better performance, the training loss was 2.2085 with an accuracy of 41.96% while the validation was 3.3143 with an accuracy of just 22.5%. The initial epoch had an accuracy of under 2%, with the increase in classes it requires a lot more epochs of training to improve the performance of the model as the weights need to be adjusted to a greater degree to have an affect on the output classes. Additionally the features between the different dog breeds can be very similar and could be challenging for a model to differentiate thus requiring more training time, or a change in the learning rate to accelerate the training by making larger changes in the weights at each epoch.

2 Experiment Two

The second experiment the saved model from the Stanford dogs dataset was used, although not required 2 variations were tested in both of them the number of epochs were increased to 50, the output layer was changed to allow for the binary classification of cats and dogs, as well as the scoring which was changed back to Binary Cross-Entropy. The first variation the weights were frozen the initial scores for the training was 0.6217 loss with an accuracy of 65.88% while the validation had an accuracy less than 50% similar starting point as that of the original. In contrast to the second model had not been frozen where the loss at the first epoch was 0.3722 and an accuracy of 83.26%, while the validation accuracy was 73.83% this is a much better starting point over the original increase of roughly 20%. The final epoch of the frozen weights showing a score of for the training as a loss of 0.0467 with an accuracy of 98.32%, while the validation had a loss of 0.1758 and an accuracy of 93.96% a slight improvement over the original cats and dogs dataset model. The second model of the non-frozen weights outperforming with a loss in training of 0.0249 and an accuracy of 99.15%, while the validation had an accuracy of 95.39% an increase of 2.4%.

3 Experiment Three

In the third experiment the output layer was changed as in experiment two as well as the first two convolution layers which was a 2-D Convolution layer with 128 neurons as well as a Separable 2-D Convolution layer with 256 neurons. Similar to the second experiment the weights are frozen on the unchanged layers in this model the starting point remains very similar while the loss in the final epoch for training is 0.0465 and accuracy of 98.28%, while the validation accuracy is 95% which is a 2% improvement over the original and 1% over the second experiment. The unfrozen model had a similar starting point to the second experiment although the validation accuracy had increased in the final epoch the model had a training loss of 0.0330 and an accuracy of 98.88%, while the validation accuracy was 94.39%. In both of these models the validation scores would vary about 10 - 15% from epoch to epoch.

4 Experiment Four

In the fourth experiment the output layer was changed as in the prior 2 experiments as well as the last two convolution layers which were both Separable 2-D Convolution layers with the second to last having 728 neurons and the last having 1024 neurons. The frozen model performing similar to that of second experiment while the final epoch the unfrozen model for training having a loss of 0.0221 an accuracy of 99.22% and a validation accuracy of 96.24% like the previous model the accuracy of the validation ranging from 10-15% in the epochs.

5 Conclusion

It would appear that transfer learning has a good effect on how well the model can perform from the initial epoch, being that the original model had fewer epochs it may be that the performance could have been improved with more time. Although the saved model from the Stanford Dogs dataset did not have a good performance it still had an impact on the

new transferred learning models. Given more epochs of the Stanford Dogs dataset model it is possible the initial epoch of the new model could have been increased further. The best model performance was the final experiment where the layers closer to the output layer its possible that end position of layers has a greater affect on the transferred models than those at the beginning.