1 Transfer Learning

We tried two ways of finetuning the model. The first one was to freeze all model parameters and only re-train the weights of the final layer. The second method of fine tuning was to retrain the whole network with a small learning rate. The idea behind re-training the whole network is that earlier layers in the network contain more generic features which is useful for detecting a wide variety of features. The latter layers are more specific to the details of the image classes.

2 Reflection on different methodologies

Re-training only the final layer of the network was faster as expected (3mins 55s). In comparison, re-training the whole network for 10 epochs with the same learning rate of 1e-4 took slightly longer (5mins 48s). However, comparing the accuracy on validation set, it is clear that retraining the whole network on our relatively small dataset caused overfitting (99.9% accuracy on validation set). By re-training only the final layer, we get better generalisation performance as shown in figure 1. This method gave final accuracy of 71.8 % on validation set.

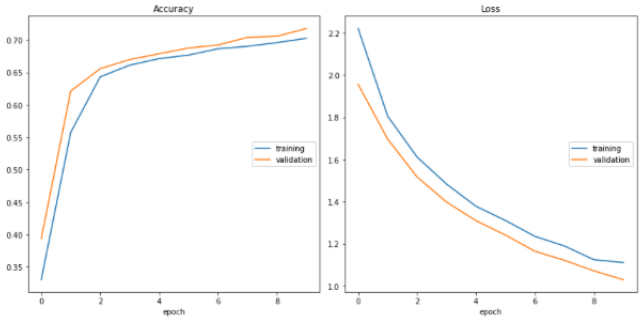

Figure 1: ResNet re-trained only on final layer gives good generalisation.

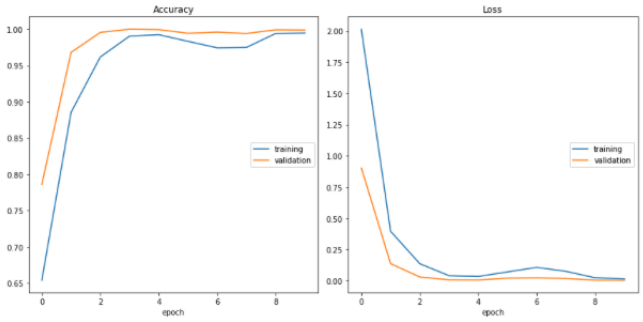

Figure 2: ResNet re-trained on all layers causing overfitting.