Assignment 1 Report
Transfer Learning on Dogs vs Cats Dataset

Arian Ghadirzadeh

April 2025

Introduction

This assignment explores the use of Convolutional Neural Networks (CNNs) and transfer learning to classify images of cats and dogs. We perform four experiments:

- Train a CNN from scratch on the Cats vs Dogs dataset (Experiment 1a)
- Train a CNN from scratch on the Stanford Dogs dataset (Experiment 1b)
- Apply transfer learning by reusing pre-trained weights and modifying output layers for Cats vs Dogs classification (Experiments 2-4)

Experiment 1a: Cats vs Dogs from Scratch

We trained a CNN on the Cats vs Dogs dataset using a learning rate of 0.0001 and 30 epochs.
Train Accuracy: ∼98%
Validation Accuracy: ∼74%

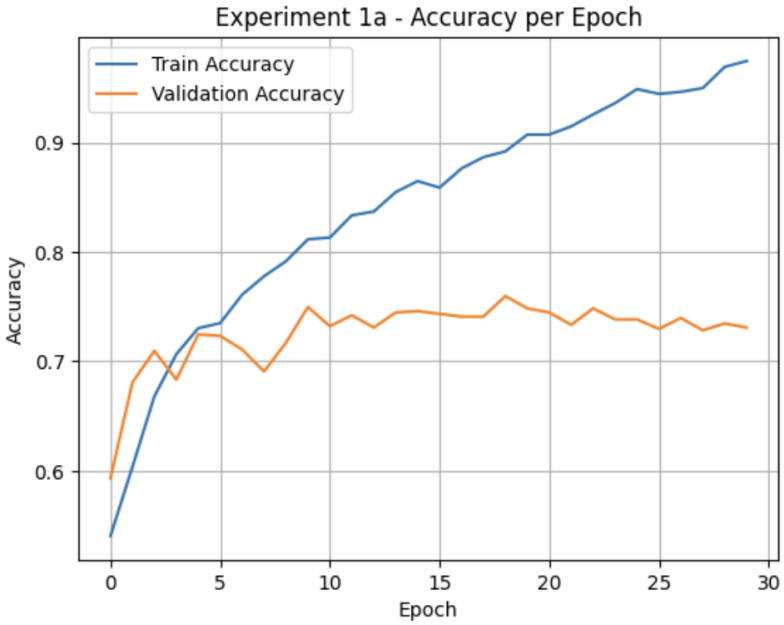

Figure 1: Experiment 1a - Accuracy per Epoch (Cats vs Dogs)

Experiment 1b: Stanford Dogs (Multiclass)

This model was trained on the Stanford Dogs dataset to classify 120 breeds. The resulting model is used as a base for Experiments 2 to 4.

Train Accuracy: ∼71%
Validation Accuracy: ∼2–3%

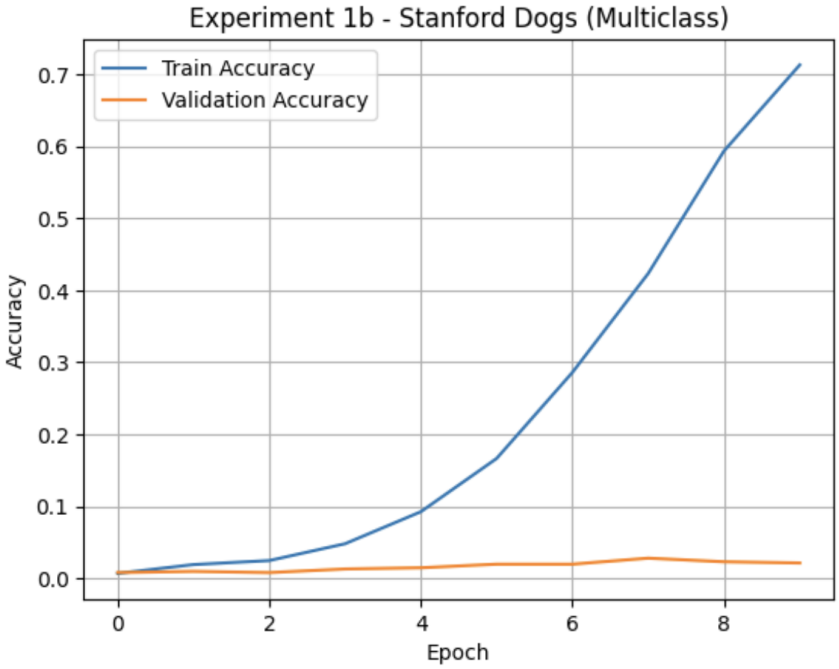

Figure 2: Experiment 1b - Accuracy per Epoch (Stanford Dogs)

Experiment 2: Replace Output Layer Only

We load the model trained in 1b and replace only the output layer for binary classification (cats vs dogs).

Train Accuracy: ∼55%
Validation Accuracy: ∼57%

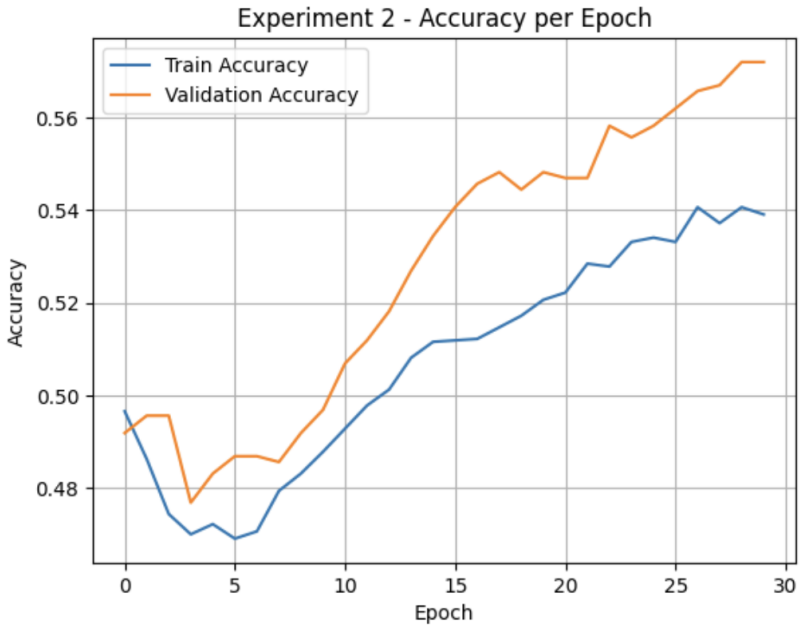

Figure 3: Experiment 2 - Accuracy per Epoch

Experiment 3: Replace Output + First Two Convs

In this experiment, we replace the output layer and the first two convolutional layers while freezing the rest.

Train Accuracy: ∼100%
Validation Accuracy: ∼69%

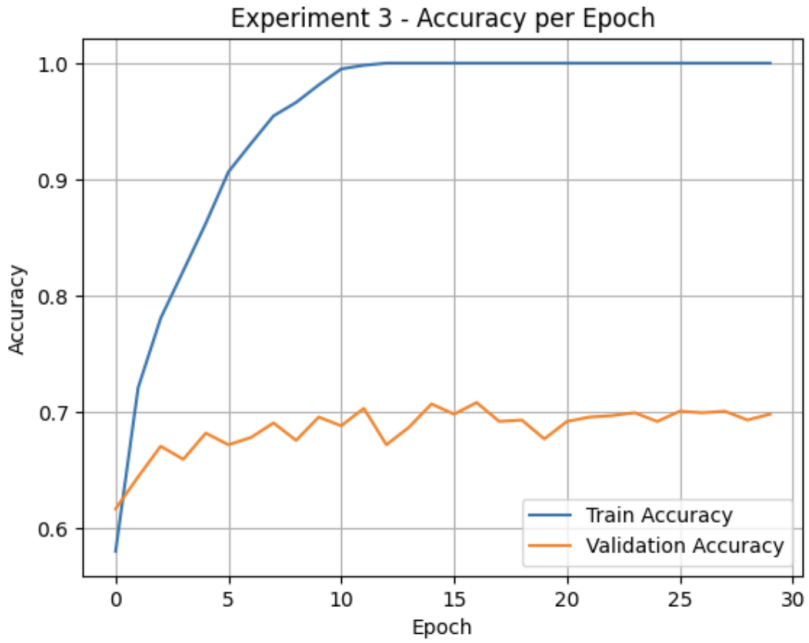

Figure 4: Experiment 3 - Accuracy per Epoch

Experiment 4: Replace Output + Last Two Convs

We now replace the output layer and the last two convolutional layers while keeping the rest frozen. This allows high-level representations to be better tuned for the new classification problem.

Train Accuracy: ∼100%
Validation Accuracy: ∼70%

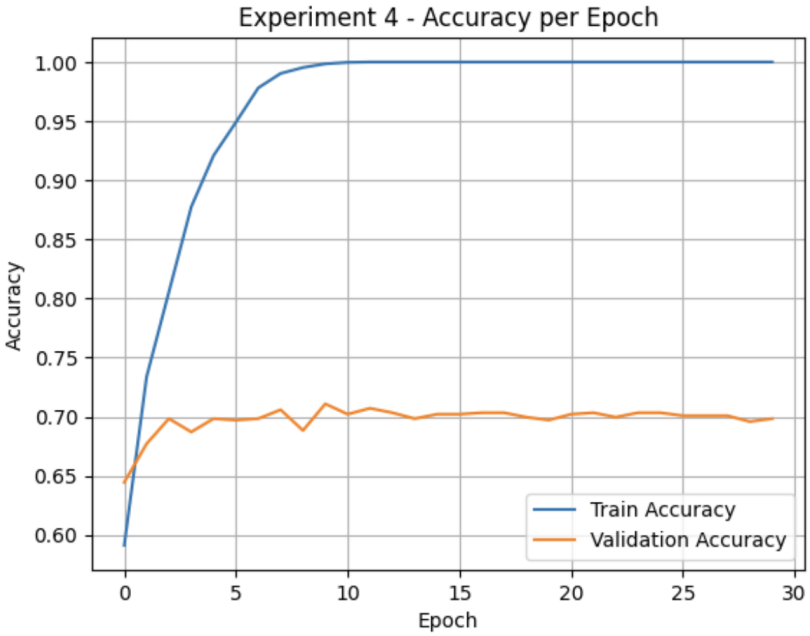

Figure 5: Experiment 4 - Accuracy per Epoch

Discussion

The baseline model trained from scratch (Experiment 1a) performed the best overall on validation accuracy, likely because it was fully optimized for the specific Cats vs Dogs dataset.

Experiment 2, which kept all pre-trained layers frozen and only replaced the output layer, performed poorly due to its limited ability to adapt features to the new task.

Experiments 3 and 4 showed a clear performance boost by fine-tuning part of the convolutional base. Notably, Experiment 4 achieved the best validation accuracy among all transfer learning settings, showing that adapting higher-level features was more beneficial.

Conclusion

The experiments show that transfer learning is effective even when the base model was trained on a different dataset (Stanford Dogs). Replacing only the output layer (Experiment 2) gave modest performance, but modifying convolutional layers improved generalization. Experiment 4 yielded the best validation accuracy among transfer learning settings, while training from scratch (Experiment 1a) remained the best overall performer.