Convolution Assignment — Cats & Dogs: Findings (Dibakar Bhowal)

Overview

This report summarizes the experiments and results from the Cats & Dogs convolution assignment. I've followed the four-step protocol: train-from-scratch experiments for three specifically sized dataset subsets (Step 1–3), then repeated the same three subsets using a pretrained VGG16 model with feature-extraction + fine-tuning (Step 4).

Methods and Optimization Techniques

Main techniques used:

- Data augmentation (RandomFlip, RandomRotation, RandomZoom)
- Rescaling / VGG preprocess for pretrained models
- Dropout (0.5) for regularization
- ModelCheckpoint to keep best validation model
- Optimizer: RMSprop (and RMSprop with very low LR 1e-5 for fine-tuning)
- Pretrained VGG16 used for transfer learning with a classifier head, then fine-tuned top layers
- Early stopping strategy implicit via saving best model by validation loss

Results (Test accuracy)

Experiment	Train / Val / Test	Test Accuracy
Scratch - Step 1	1000 / 500 / 500	0.7580
Scratch - Step 2	2000 / 500 / 500	0.8400
Scratch - Step 3 (ideal)	2000 / 1000 / 1000	0.8250
Pretrained - Step 1	1000 / 500 / 500	0.9660
Pretrained - Step 2	2000 / 500 / 500	0.9740
Pretrained - Step 3 (ideal)	2000 / 1000 / 1000	0.9830

Analysis & Insights

1. Transfer learning (pretrained VGG16 + fine-tuning) outperformed the model trained from scratch by a large margin for all sample sizes. Even with only 1,000 training images, the pretrained model reached ~96.6% test accuracy, whereas the scratch model reached ~75.8%.

2. For the scratch models, increasing training data from 1,000 → 2,000 improved test accuracy substantially (0.758 → 0.840). The 'ideal' run used the same training amount as

Step 2 (2,000) but larger validation/test splits; test accuracy for that run was slightly lower (0.825), likely because the larger test set is a stricter estimate and/or because of variance in which images were selected.

3. Pretrained models improved significantly with more train data (96.6% → 97.4% → 98.3%), indicating diminishing returns but still measurable gains from additional labeled data when combined with transfer learning and fine-tuning.

4. Practical takeaway: if compute or labeled data are limited, using a pretrained network with augmentation + fine-tuning is recommended. If training from scratch, increasing labeled training data helps a lot but still lags behind transfer learning in this experiment.

Recommendations

• Using a pretrained model (VGG16 here) and fine-tuning top layers produces best results under modest dataset sizes.
• If training from scratch is required, collecting substantially more labeled images (>>2000) and use of regularization and augmentation is essential.
• Keeping a held-out test set (>=500) for stable evaluation and using ModelCheckpoint to avoid selecting overfit checkpoints is beneficial.