Cats vs. Dogs Image Classification Using Deep Learning

Pasnoor Prahasith

Abstract

This project focuses on classifying images of cats and dogs using deep learning techniques. We implemented a Convolutional Neural Network (CNN) with a pre-trained ResNet18 model to achieve binary classification. The dataset consisted of labeled images of cats and dogs, and we applied various data transformations to enhance model performance.

1 Introduction

In recent years, image classification has become a crucial task in computer vision, with applications ranging from object recognition to autonomous driving. This project aims to build a classifier that distinguishes between images of cats and dogs using a deep learning approach. We utilized a pre-trained ResNet18 model to leverage transfer learning, which has shown promising results in similar tasks.

2 Methodology

2.1 Dataset

The dataset used for this project consists of images of cats and dogs. The images are stored in a directory structure where the training images are organized into ‘cat‘ and ‘dog‘ subdirectories. The test images are similarly organized.

2.2 Data Preprocessing

To prepare the data for training, we applied the following transformations:

- **Training Data:**
  - Random Resized Crop

  - Random Horizontal Flip
  - Normalization

- **Test Data:**
  - Resize
  - Center Crop
  - Normalization

2.3 Model

We used the ResNet18 model, pre-trained on the ImageNet dataset, as the backbone for our classifier. The final fully connected layer was modified to output two classes: cat and dog.

2.4 Training

The model was trained using the following settings:

- **Optimizer:** Adam
- **Learning Rate:** 0.001
- **Loss Function:** Cross Entropy Loss
- **Epochs:** 15

3 Results

The model was evaluated on a test set, and the training and test accuracies over epochs are as follows:

The best validation accuracy achieved was 0.9723.

4 Conclusion

The project successfully implemented a binary image classifier for cats and dogs using a pre-trained ResNet18 model. The model achieved an accuracy of 97.23% on the test set. This demonstrates the effectiveness of transfer learning and fine-tuning on pre-trained models for image classification tasks. Future work may involve exploring other architectures, adjusting hyperparameters, or expanding the dataset to include additional classes.

Epoch	Train Accuracy	Test Accuracy
1	0.8578	0.9181
2	0.8892	0.9524
3	0.9005	0.9522
4	0.9011	0.9325
5	0.9127	0.9494
6	0.9161	0.9625
7	0.9180	0.9580
8	0.9207	0.9566
9	0.9237	0.9410
10	0.9276	0.9705
11	0.9287	0.9661
12	0.9292	0.9683
13	0.9334	0.9679
14	0.9339	0.9673
15	0.9338	0.9723

Table 1: Training and Test Accuracy Over Epochs

5 References

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- PyTorch Documentation. (2024). *https://pytorch.org/docs/stable/index.html*
- torchvision Documentation. (2024). *https://pytorch.org/vision/stable/index.html*