HW1 Fine-Grained Image Classification

Outline

- Introduction
- Methodology
- Experiment
- Code and Reference

Introduction

- The purpose of this project is to classify 196 different categories of cars with different brands and types.
- Training dataset: 11185 images with ground truth
- Testing dataset: 5000 images without ground truth

Methodology

- Data pre-process and augmentation:
  a. Resize images to (512, 512)
  b. Random crop images to (448, 448)
  c. Random horizontal flip
  d. Color jitter (brightness: 0.126, saturation: 0.5)

- Model Architecture

  Resnet50 provided by torchvision library with pretrained weights. The output dimension should be modified to 196 to fit the number of the categories of the dataset.

Methodology (cont.)

- Hyperparameters
  - Epoch: 120
  - Batch size: 64
  - NN weight initialize: kaiming normal
  - NN input image size: 448 * 448
  - Loss function: cross entropy
  - Optimizer: stochastic gradient descent with momentum=0.9
  - Learning rate: Start with 0.01, multiply by 0.99 every epoch

Experiment

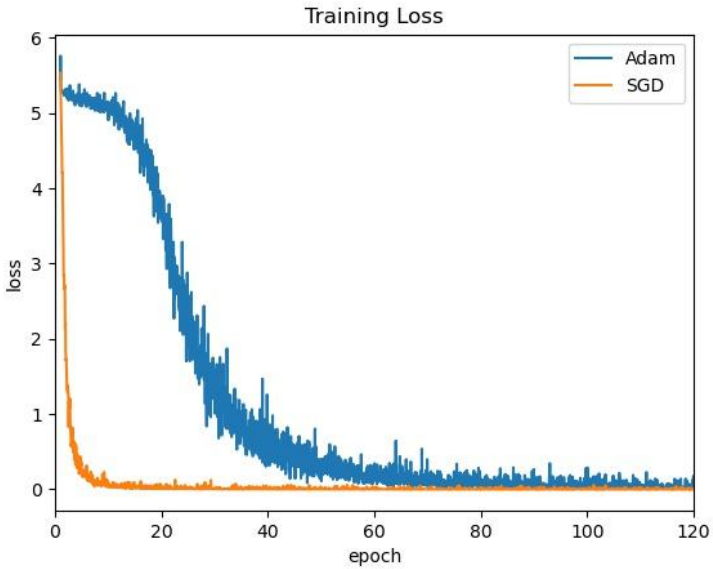
- Hardware information
  - CPU: i9-10900X
  - GPU: RTX 2080ti * 3
  - RAM: 62G

- Training time for 120 epochs: about 6hrs

Experiment (cont.)

- I've tried different optimizer supported in pytorch and found that SGD with momentum performs better than Adam in both the decreasing speed of training loss and the testing accuracy.

- Testing accuracy after 120 epochs:
  - SGD: 0.948
  - Adam: 0.823



Code and Reference

- Github Link

- References are listed in README.md in github repository.