Selected Topics in Visual Recognition using Deep Learning

Homework 1

陳泊均 0616043

Project repository: https://github.com/SafiJ/Visual-Recognition-using-DL

Introduction

The challenge is to classify car images to their categories (consist of car brand & produced year) correctly, using the given dataset (including training_data & testing _data) provided by TA. There're 11185 and 5000 images in training_data and testing_data respectively containing 196 classes in total, and all these data are from Stanford Car Dataset.

Data Processing

At first, I resized each image to 224x224, but the accuracy is quite low (about 84%), so I tried to raise the resolution of resized images which is 400x400 (too large would cause CUDA out of memory). Besides, I also added some transformation in those images, such as HorizontalFlip, RandomRotation etc.

Model Training

I decided to use Resnet as training model. After trying resnet34 and resnet50, I decided to use resnet34 as my car classifier model, the accuracy keeps rising when adding epochs as the picture showed below. Finally I get 91.86% by resnet34 training after 30 epochs. But when epoch exceed 30 it seems to be overfitting.

On the other hand, the highest accuracy of using resnet50 only goes up to 90.8%.
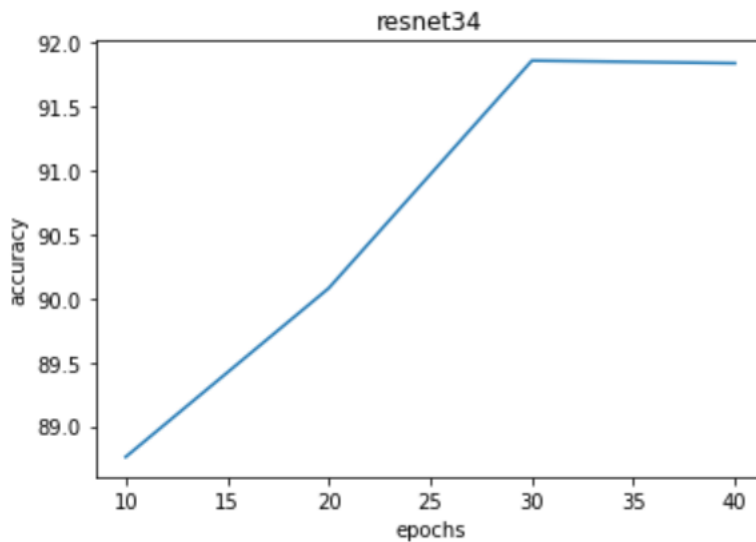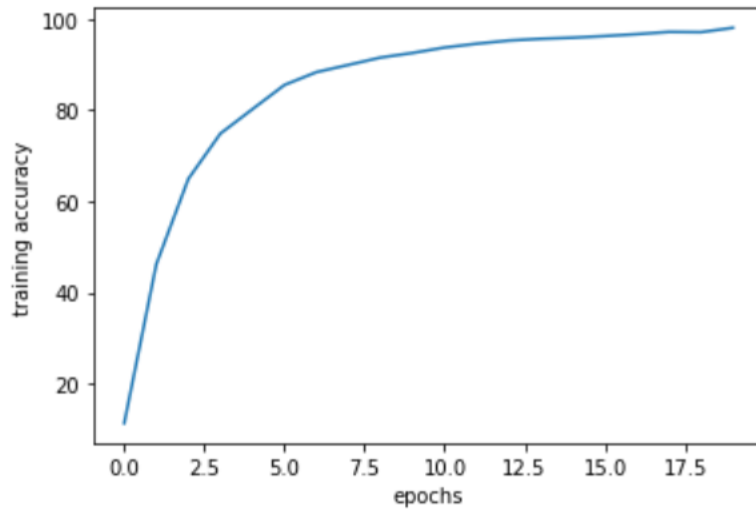
resnet34

Summary

The results between resnet34 and resnet50 are pretty much the same. The main reason that affects the accuracy is the image size (larger image size gets higher accuracy). I guess it's because there's more features can be extracted by higher image resolution.

Reference

https://github.com/wengsengh/Car-Models-Classifier/blob/master/car_models_classifier.ipynb

https://www.kaggle.com/deepbear/pytorch-car-classifier-90-accuracy/notebook#Model-training-function