Machine Learning Engineer Nanodegree

Capstone Proposal

Joel Haas
September 15, 2019 (Review notes: https://review.udacity.com/#!/reviews/1971290)

Proposal

Domain Background

This is an object detection, classification and recognition project. Large areas of research include object detection, classification, and recognition. Object detection is processing an image and detecting discreet objects within that image. Object classification is taking the objects detected and identifying what those objects are. And finally, object recognition is taking those classified objects and recognizing the differences between them. Object recognition is the most difficult of the three mentioned, but nevertheless, there is a lot of research in these areas.

It is this technology that enables companies like Apple to use face recognition to unlock a user's iPhone. The iPhone detects a face, classifies it as a human face, and then recognizes that the human face is the owner of the phone. It's an exciting area of research that I would like to improve in.

Problem Statement

My project is an object detection, classification and recognition project. Given an image as input, I want to detect whether a car is in the image or not, classify the car as a corvette or not, and then if it is a corvette, recognize the specific generation of corvette – C1 through C8.

Datasets and Inputs

The data for this project consists of thousands of images. There need to be thousands of images without cars and thousands of images for each generation of corvette.

I will need to scrape the web for enough images to train on. This is one risk to the project – whether I can scrape enough images for the algorithm to learn on.

For each generation of corvette scraped, I will need thousands of images of different angles of the car. This is another risk to the project – identifying a specific generation corvette when only an aerial picture is provided or when the picture is just of the back of the corvette. This is much harder than if all the corvette pictures were from a broad side of the car.

Each of the images will need to be classified accordingly. Images with and without cars. Images with and without corvettes. And corvette images classified by generation – C1 through C8.

Solution Statement

I will leverage transfer learning to train the algorithm. By leveraging a CNN that has already been pre-trained, I will get a head start on classifying whether a car is in the picture or not. The solution will modify the pre-trained model and add additional layers (as time and compute allows). I will measure my progress by calculating the accuracy of my predictions.

Benchmark Model

I will use the Dog Breed Classification project as a benchmark and the accuracy of predicting dog breeds. Predicting the breed of a dog is like predicting the generation of corvette (C1 through C8). After several google searches, it looks like 80% accuracy is the benchmark for the dog breed classification project, so I think that is also a good target to shoot for when predicting the corvette generation.

Evaluation Metrics

I will use accuracy as the evaluation metric to quantify the performance of both the benchmark model and the solution model. I'm trying to predict, with accuracy, whether an image has a car or not, whether that car is a corvette or not, and what generation of corvette it is.

Project Design

For this project, the following workflow will be implemented: