DSC 672: Predictive Analytics Capstone

Milestone 1 - Project Proposal

Stanford Car Dataset: Vehicle Recognition Project

Alexandre Girault (Team Manager)
Chris Shaffer
Sean Sungil Kim
Ryoh Shinohara
Yanxi Cai

Spring 2019

I. Data Description

The public Stanford Cars Dataset contains total 16,185 images of cars. There are a total of 196 classes of cars in this dataset. The data is split in half to be used as training and testing sets. The data also comes with class labels and bounding boxes for all images. The classes are typically at the level of Year, Make and Model (e.g. 2012 Tesla Model S or 2012 BMW M3 coupe). The sizes of each image are different. Utilization of the bounding boxes is essential in the pre-processing phase to first obtain images that focus on the objects of interest, which in this case are the vehicles. The actual images are in JPG format, but the data comes zipped in TGZ/TAR format.

Data source: http://ai.stanford.edu/~jkrause/cars/car_dataset.html.

II. Project Description

The Stanford Car Dataset will be utilized to build a vehicle recognition predictive model. The ultimate goal of the model is to classify a car's year, make and model given an input image. This model could be further developed to be used in creating a mobile application that assists users in identifying cars of interest. The users would simply take a picture of the vehicle of interest and the application would return information (year, make and model) regarding the recognized vehicle. The users could also input a picture found on the Internet. Partnerships with other car dealership websites could be beneficial in enhancing the application quality, since the recognized vehicle name would be used in searching the partners' database to obtain valuable information such as availability, price and so on. An improved model would result in direct reviews/subscription profit. This application could help people who are not familiar with cars or who simply want quick information without searching the Internet themselves.

Another potential development idea of this project would be for traffic law enforcement. Traffic AI is a huge market globally. One example of this would be the China Transinfo Technology Corp. They focus on extracting features of vehicles the moment they appear in security cameras, which can help the police to track the targeted cars.

Different classification algorithms such as RandomForest, SVM, some of the boosting approaches, in addition to Convolution Neural Networks will be explored. Some of the feature extraction and feature selection methods will be explored as well, for the non-deep learning classifiers. The values of these models will be quantified in terms of performance and cost. Furthermore, in order to evaluate the true performance of the model, 30 images or more will be added to the validation set to gauge the true real-world predictive power.

Possible research questions:

- What level of information can be obtained about vehicles through image recognition algorithms? Would classifying only the Make and Model result in higher performance scores?
- What would the ideal reshaped size of the input images be in terms of performance and cost?
- What kind of feature extraction methods will have the most impact in increasing the performance and/or decreasing cost?
- What would be the best way to balance both performance and cost?
- For CNNs, what kind of ConvNet architecture will be the most valuable in terms of predictive power and cost?
- Does the accuracy of CNN models outweigh the benefit of non-NN classifier's runtime speed?