Machine Learning Engineer Nanodegree Udacity

Dibakar Sigdel

May 27, 2017

Capstone Proposal

Domain Background

Object Detection [1] is one of the exciting part of the study in Computer Vision and Machine Learning. For example human detection [2] in surveillance cameras, face and facial expression detection[3] in speech videos etc. This project is greatly inspired by 'Vehicle Detection and Tracking Project' of the **Self Driving Car Nanodegree** which is offered at Udacity. The main focus on this project is to build a model for image classification and using it to detect and track the identified image in the videos. More specifically, it will detect and track vehicle in the video.

Problem Statement

The statement of the problem is **Vehicle detection and tracking in a video using classifier trained over vehicle and non vehicle image data set**. The solution to the problem is that, it should be able to select

proper pixels in the image frame in videos as vehicle and draw a rectangular boundary. More technically, the accuracy of the classifier will be calculated with respect to train, test data as well as the performance of classifier on videos. The results thus obtained should be replicable using same data set.

Data Set and Inputs

This project will build a classifier using two sets of data called vehicle and non-vehicle. These data and test video is obtained from the link provided by Udacity at 'Vehicle Detection and Tracking project' of 'Self Driving Car Nanodegree' program.

The data set includes 8792 cars and 8968 non-cars. Each image is of size 64 by 64 with three color chanels i.e. of dimension: (64, 64, 3). Below is few sample images from the data set.
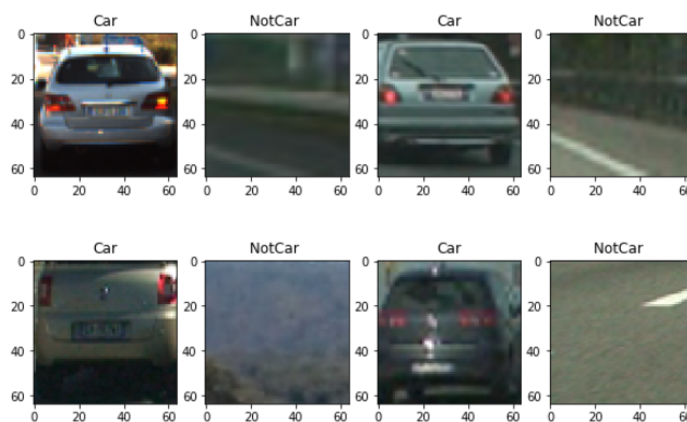

Figure 1: car and not-car samples

Once a model is ready, it will be used to test it's performance on test video. Following are the data sources:
1. Vehicle Images: click here
2. Non-vehicle Images: click here
3. Test Video: click here

Solution Statement

The problem of detecting vehicle in the test video will be solved by using techniques of Computer Vision for feature extraction and and Machine Learning for search of best classifier model. Feature extraction steps will include use of the concept of Histograms of Oriented Gradients HOG, and color features like color histogeam and spatial distribution of pixels.

Different schemes and hyperparameters for SVM classifiers for eample *linear, rbf, 'C'* will be tried. For parameter tuning, scikit-learn's *GridSearchCV, RandomizedSearchCV* will be used. Best classifier model will be extracted using labeled data set and it will be used to the test video. Once the part of the image representing vehicle is identified in each frame of video, a rectangular box will be drawn around it for vehicle (as shown in figure below) tracking purpose. The correctly counted vehicle in the duration of the video will count the success of the algorithm.

Figure 2: Rectangular box around detected vehicle

Benchmark model

This project will use report on Vehicle Detection in an Image and A Two-Stage Approach to People and Vehicle Detection With HOG-Based SVM as our bench mark model. These models use the concept of Histograms of Oriented Gradients HOG, for feature extraction purpose and some other geometrical aspects like edge detection etc. as well as SVM for classifier. Our aim in this project is also to extract color and HOG feature before we train the proper classifier.

Evaluation Metric

Two different metrics will be used in this project.

- First one includes the accuracy of the trained support vector machine on training and test data.

- The second includes the count of correctly detected vehicles and false positive cases. The actual number of the cars will be counted manually and it will be compaired to the vehicle detected by the algorithhm. Since we are using a binary classifier and we want to account for false-positives we will use the following formula to calculate accuracy:

$$\text{Accuracy} = \frac{(\text{true positives} + \text{true negatives})}{\text{size of dataset}} \tag{1}$$

  Here are true positives are cars identified correctly. True negatives are image segments without a car identified as such. The size of the dataset is the total number of image segments classified. The detection where rectangular box is off the actual vehicle image will not be counted as true positive.

Project Design

The whole project will be divided into following steps:

1. **Getting Familiar with data and exploratory data analysis** : On this step, a gentle study of data will be done to be familiar with it and also a exploratory data analysis will be presented before the actual feature extraction.

2. **Feature Extraction** : This step will provide detail of feature extraction. Proper image size will be selected. Various color chanels like: 'BGR','HSV','LUV','HLS','YUV','YCrCb' will be studied for better feature. HOG features and color features are the main part of the features.

3. **Train a linear SVM**: This step will provide detail of training process, selection of parameters and theory of the model. Feature extracted from step 2 will be normalized and splited into train test part and will be fed in to the training process.

4. **Technique of Slidding Window Search** : This step will discuss about the utilities used for search of vehicle in the different parts of an image. An image is segemented in to a smaller parts and some parts are given priority for searching vehicle. This results in a set of windows from which suspected part of the image is extracted and tested for presence of vehicle.

5. **Search and Classifiy an image** : This step uses all steps mentioned above, pretrained classifier model uses tiny section of the large image one-by-one through sliding window technique. If the classifier predict that tiny section as a car, a rectangle will be drawn around that section.

6. **Search and Classifiy a Video**: This is the final part of the project which prepares the video along with vehicle inside a rectangular box. Since video is a time series of many images, we use the step 5 in all images in video and reprocess them to form final output video. To be more precise in the result, we will use heatmap technique in vehicle detection to catch false positive cases as well.

7. **Results and Conclusion** : This step will discuss the results and accuracy obtained in the model used above and compere it with other models available.

References

1. Object Detection Wikipedia
2. Human detection from images and videos: A survey
3. Machine Learinig Technique for Face Analaysis
4. Vehicle Detection in an Image
5. A Two-Stage Approach to People and VehicleDetection With HOG-Based SVM
6. Histograms of Oriented Gradients for Human Detection