CS 596

Project Report

Vehicle Detection using Support Vector Machine

Nahye Park

1. Abstract

Vehicle detection is a part of object detection. Generally, vehicle detection and tracking applications are used in self-driving cars, highway traffic surveillance control and urban traffic planning. It can be done with machine learning algorithms such as Deep Learning, Support Vector Machine, Decision Tree and so on. Since it's using data from images, the data of features in images is needed before training the classifier. Histogram of Oriented Gradient(HOG), Spatial Binning or Color Histogram can describe the image features. For this project, I used support vector machine algorithms for training and histogram of oriented gradient for feature extraction.

Histogram of Oriented Gradient is one of the methods that extract the features from the image. This method is used for computer vision and image processing for object detection. This is done by calculating the gradient of blocks of a fixed size and putting the calculated data into one dimension array by counting the frequency at a specific angle. For this project, a total 900 features were extracted from each image and fed into the support vector machine classifier.

Support Vector Machine is one of the supervised learning models. This method builds the hyperplane that divides the data in the high dimension space. The margin determines how far the hyperplane will be from the data. Hard margin and soft margin were considered for this project. Usually, hard margin is used for the linearly separable data set. There is a C value that is a hypermeter to control error. Low C value gives us a lower error rate and a large c gives us a higher error rate. I wanted a more flexible and exact separation, so I set the C value to 0.1 in this project.

14,208 of training data and 3,552 of test data were used in the project. As the result of the first try, images classified as false positives included the strong two vertical lines like a road. Thus, the accuracy was 0.9654 at first time and it became 0.9685 by changing settings of the project, by increasing the proportion of horizontal lines in features. False positive rate also decreased from 0.037 to 0.031. To increase the accuracy and decrease the false positive rate, additional feature extraction methods are considered for vehicle detection using the support vector machine.

2. Introduction

In recent years, people's interest in self-driving cars has soared. Self-driving cars have already come into the world, but they are not perfect yet. Accordingly, technological development is also necessary, and object detection is one of the technologies that must be developed more precisely. Object detection is useful not only in self-driving cars but also in various traffic industries like traffic management. The object detection can be done by various machine learning like Support Vector Machine, Decision Tree and Deep Learning. I'm going to analyze and evaluate the vehicle detection, which is one of the object detection, using a Support Vector Machine.

3. Literature Review

Selection of Histograms of Oriented Gradients Features for Pedestrian Detection

In this paper, they explain pedestrian detection using a linear support vector machine and histograms of oriented gradients. The author is especially focusing on the preprocessing of image data. They applied Principal Component Analysis (PCA) to feature vectors from histograms of oriented gradients. Also, they talk about the Difference of Gaussian (DOG) to get the silhouette of humans and the effective regions in the image that include edges. Accordingly, I also attempted to apply PCA and DOG to this project. However, when I applied them to this project, the execution time has been accelerated, but as a result, the accuracy has decreased. Since the goal of this project is to be more accurate than time, I decided not to apply PCA and DOG.

Vehicle Detection and Tracking Using Machine Learning Techniques

In this paper, the support vector machine is better for vehicle detection than the decision tree because accuracy of the support vector machine classifier is about 98%, while accuracy of decision tree classifier is about 98%. The accuracy of the support vector machine in this project is about 97%. Compared to the results of this project only with the results of the support vector machine, there is a difference of about 2% in accuracy. This difference seems to be due to the difference in the preprocessing process. In addition to HOG, they used Spatial Binning and Color Histogram in the preprocessing.

Support Vector Machine as a Binary Classifier for Automated Object Detection in Remotely Sensed Data

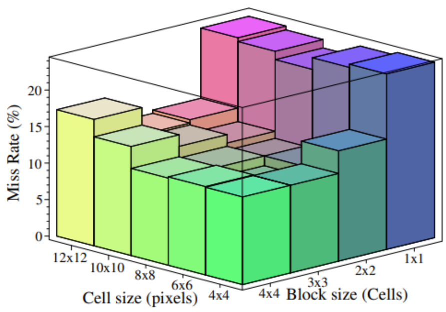
In this paper, the author classifies and detects two classes, an object class and a background, from images from remotely sensed data, rather than detecting a predetermined object. As it does not detect a specified object, the shape or the edge of the object is not important data unlike this project. That's why the author constructs the hyperplane based on the RGB pattern of the image. In vehicle detection using a support vector machine, we can add the color components into feature vectors, however, it is difficult to train the classifier successfully only with the color features.

Object Detection Using Support Vector Machine and Convolutional Neural Network - A Survey

In this paper, as a result of the experiment, Convolution Neural Network (CNN) for object detection has better accuracy than Support Vector Machine. Furthermore, object detection using the convolution neural network has a lower error rate than using the support vector machine. However, fine-tuning the supervised dataset using the convolution neural network seems important. When we look at the results of the above research paper, we can see that the Convolution Neural Network is also more suitable for object detection than the decision tree.

Histograms of Oriented Gradients for Human Detection

In this paper, lots of preprocessing methods are mentioned including gradient computation, color normalization and spatial/orientation binning. The interesting fact here is about the descriptor blocks. As we can see at figure 3.1, the combination of the smallest block size and the smallest cell size produced the largest miss rate. This means that dividing the image into tiny and small pieces does not increase the accuracy. Although the graph shows that the best option for block size and cell size is the combination of 3x3 block size and 6x6 cell size, I used 10 x 10 as cell size and 2 x 2 as block size because the smaller the number of pixels per block or cell, the longer the execution time.



[Figure 3.1 : From 'Histograms of Oriented Gradients for Human Detection']

4. Datasets

Datasets for this project are from Vehicle detection and tracking : Udacity's Self-driving Car Nanodegree. There are vehicle and non-vehicle images. The vehicle images include the right, left, middle and far side of cars and non-vehicle images include parts of scenes that can be seen while driving, such as mountains, trees, roads and sky. There are 8792 vehicle images and 8968 non-vehicle images with size 64x64. I combined the vehicle images and non-vehicle images and divided the training set and the test set into a ratio of 80 to 20. Therefore, a total of 14,208 images were used for training the classifier and 3,552 of images were used for evaluating the trained classifier.

[Figure 4.1 : Examples of vehicle images]

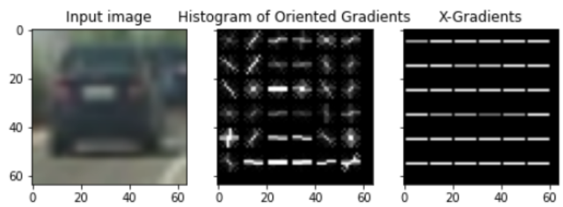
[Figure 4.2 : Examples of non-vehicle images]

5. Feature Extraction

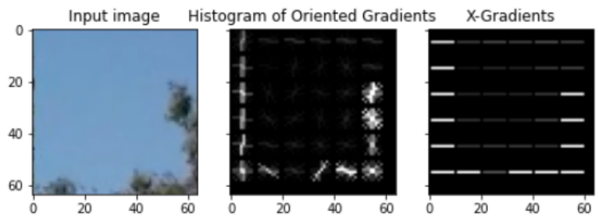
For the histogram of oriented gradients, the orientations are considered as 8. For pixels per cell parameter, I chose 8 x 8 at the first try, but it takes a long execution time and too much memory for training the classifier. Thus, I fixed the pixels per cell as 10 x 10 later. Cells per block were 2 x 2 in the project settings so it's 20 x 20 pixels in this case.

When evaluating the classifier, the images misclassified as a vehicle contain strong vertical lines. So I added the features extracted using 2 orientation once more to the feature array that will be fed into the classifier. As a result, it increased the accuracy of the classifier. According to this, a total 800 features were fed into the support vector classifier with these options. Figure 5.1 and figure 5.2 show the direction and magnitude of gradient of vehicle image and non-vehicle image.

I considered the Difference of Gaussian (DOG) for the preprocessing before extracting the features from the image and Principal Component Analysis (PCA) for selecting the features from histogram of oriented gradient. However, after the experiments, both options actually reduced the accuracy of the classifier in this project. Thus, I didn't apply them to the project.



[Figure 5.1 : Gradients of vehicle image]



[Figure 5.2 : Gradients of non-vehicle image]

6. Support Vector Machine (SVM)

Support vector machine classifier constructs the hyperplane in high dimension space. The dimension is determined by the number of features fed into the classifier. Thus, when we have dataset

$$D = \left\{ \left(x_i, y_i \right) \mid x_i \in \mathbb{R}^p \, , y_i \in \{-1, 1\} \right\}_{i=1}^{n}$$

where n is the number of features and p is the dimension. The goal of a linear support vector machine is to design the best hyperplane that classifies training vectors in two classes.

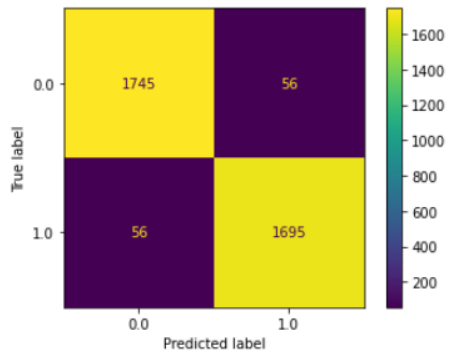
$$g(\vec{x}) = w^T \vec{x} + b$$

w is a weight vector and b is a linear coefficient. The result of classification in this project will be

$$predicted\ y = \begin{cases} vehicle\ class,\ if\ \ g(\vec{x}) \geq 1 \\ non-vehicle\ class,\ if\ \ g(\vec{x}) \leq -1 \end{cases}$$

7. Result

According to evaluating the classifier with 3,552 of test data, accuracy is 0.9685. It was 0.9654 when using the default value, which is 1, for margin in the support vector machine classifier. Also, accuracy has been increased by adding features from 2 orientations of histogram of oriented gradient twice. As the accuracy increased, the false positive rate also decreased.

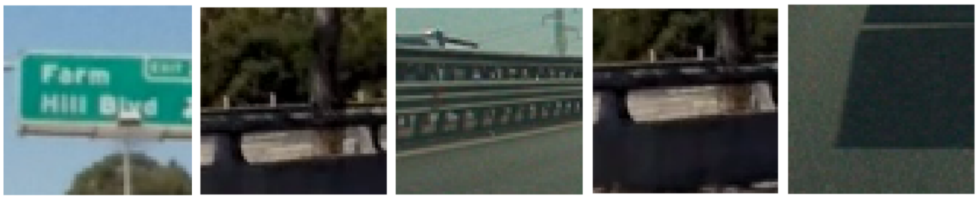


[Figure 7.1: Visualization of confusion matrix]

```
true negative 1745
false positive 56
false negative 56
true positive 1695
False Positive Rate = 0.031093836757357024
```

[Figure 7.2: Result and false positive rate]

Figure 7.1 shows the confusion matrix of the result and figure 7.3 shows some cases of misclassification as a vehicle, which are false positive cases. It includes the sky, trees and roads. So I tried to apply the color threshold for checking the intensity of the color blue, green and black but it didn't affect the false positive rate and accuracy of the classifier.



[Figure 7.3: Examples of false positive images]

8. Conclusion

According to evaluation of this project, vehicle detection using a support vector machine with histogram of oriented gradients is good. Using a linear support vector machine is good to handle lots of features from histogram of oriented gradients. When we look at the last result of false positive, figure 7.3, a rectangle shape is in most of the false positive images. So we can consider other additional feature extraction methods that are not related to the shape of edges, such as colour histogram. The experiments with other feature extraction methods would be needed for improving vehicle detection using a support vector machine.

9. References

Mustafa, S., 2019. *VEHICLE DETECTION AND TRACKING USING MACHINE LEARNING TECHNIQUES*.

P. D. Wardaya. 2014. Support Vector Machine as a binary classifier for automated object detection in remotely sensed data. *IOP Conference Series: Earth and Environmental Science*, vol. 18.

R. Khosla, Y. Singh, and T. Balachander. Object detection using support vector machine and Convolutional Neural Network - A Survey. *International Journal of Engineering & Technology*. vol. 7. no. 2.24. pp. 428–430. 2018.

T. Kobayashi, A. Hidaka, and T. Kurita. Selection of histograms of oriented gradients features for pedestrian detection. *Neural Information Processing*, pp. 598–607.

N. Dalal and B. Triggs. 2005. Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.