The Vehicle Detection Based on Linear SVM

Hongyi Luo
Computer Science of Engineering
UNSW
Sydney, NSW, Australia
z5241868@ad.unsw.edu.au

***Abstract*—Computer Vision (CV for short) is a current hot topic in artificial intelligence. This report is going to introduce which kind of CV methods can use to solve vehicle identification problems. Also, how to use these traditional CV methods to implement it.**

***Keywords*—computer vision, supervised learning, vehicle identification, support vector machine**

I. Introduction And Background

Nowadays, with the development of artificial intelligence technology, computer vision has become an indispensable part of life. This assignment requires us to complete the identification of the vehicle based on the provided data set. This task is still full of challenges these days. How to train an efficient and accurate learning model is the main focus of the task. Some researchers have achieved object detection models through convolutional neural networks. Vehicle detection is a two-classes classification problem, which is also supervised learning [1] Facing the problem of how to implement object detection, we must start with the type and structure of the data. There are three parts of datasets which are training data, test data, and support data in this project. Without using deep learning methods, HOG feature extraction and colour histogram [2] can be used to do the feature extraction. Then using the linear SVM [3] to train the vehicle detection model for this assignment. After training, using the heat map and window sliding methods to separate the image. Then using the training model to do the detection.

II. Methods And Implementation

A. Data Processing

Before doing the training, the first step is to process the data. In this assignment, selecting the supporting data as the training data is enough to train the classifier. Based on the supporting data, cutting the image based on the annotation file. By using the location data from the JSON file, it can finally cut the vehicle images from the original images. These vehicle images are also called the region of interest, for short is ROI. Also, using the random number to generate the non-vehicle ROI. After doing that, the training data of vehicle ROI and non-vehicle ROI has been gotten.

B. Feature Extraction

In this project, choosing the histogram of oriented gradients as the descriptor of the feature extraction. This first advantage is that HOG feature extraction has benefits on edge detecting and texture detecting [5]. Furthermore, doing the vector concatenate is easier to implement.

The formula (1), which is showing above, this G means the value of gradient, also the θ is the orientation of the gradient.

$$\begin{cases} \nabla G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \\ \theta(x,y) = tan^{-1}\frac{G_y(x,y)}{G_x(x,y)} \end{cases} \quad (1)$$

From the formula (1), it is not difficult to understand that HOG feature extraction [2] processes each cell in the images, while getting the gradient and orientation. It maintains good invariance to image geometric and optical deformations. Considering the influence of light [2] and other factors in the vehicle detection task, HOG feature extraction can not only eliminate the influence of sunlight but also can reduce the dimension of datasets.

C. Linear Support Vector Machine

The support vector machine is a traditional supervised learning model [3], which is also a good approach for classification issues. SVM performs well in small sample machine learning problems. Furthermore, images also have high-dimensional data sets, selecting SVM which has a strong generalization ability is a good choice. Here is Fig. 1. of the support vector machine.

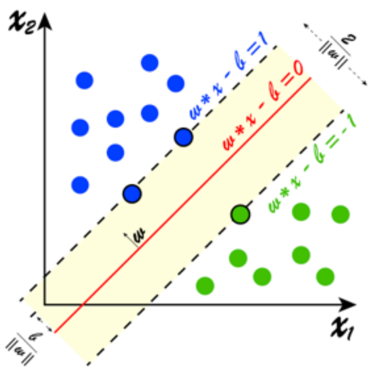

Fig. 1. SVM model

D. Heat map and window sliding

The heat map is used to display the correlation coefficient matrix of a set of variables. Through the heat map, observing the difference in the magnitude of the value has become very intuitive. In this project, using the heat map can find the changes of feature easily. Furthermore, it is convenient to set up a sliding window. Although the sliding window is a dynamic method, Although the sliding window is a dynamic algorithm, after the start points and endpoints are determined, the range of the window is also fixed. This is convenient to implement the movement of sliding windows by control the start and end points.

III. EXPERIMENT

A. Data Processing Part

The first point is to ignore data whose length and width are less than 30 when the JSON tag is obtained. It is difficult to get a feature from those images which size is less than 30*30. Besides, ignoring these data can greatly reduce the amount of data and improve data efficiency.

The second point, to make the generated positive and negative sample data consistent, when acquiring the positive sample data, collect the negative sample of the current picture. Resizing the vehicle ROI which is the positive data as 64*64. When the positive sample data is obtained, change its size to 64*64. At the same time, use the Python random number tool to generate negative samples of size 64*64.

When the positive sample data is read, it can be found that the amount of positive and negative sample data is consistent as well as the data size. The data volume of the positive and negative samples is the same, both are 10713.

Here is the sample Fig. 2. of vehicles ROI and non-vehicles ROI.

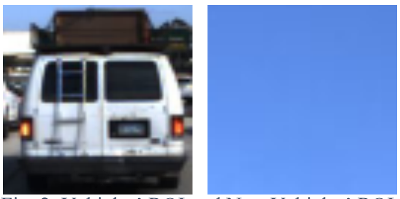

Fig. 2. Vehicles' ROI and Non-Vehicles' ROI

B. Feature Extraction

Hog Feature Extraction

When using hog for feature extraction, pay attention to these parameters. The first parameter is called 'visualize', which is used to display the HOG image and set as False. Other parameters are set to default values. Finding more information on the sci-kit image.

Original Pixel Information

Doing the extraction of the Original image colour information by using the bin spatial function is an important step to collect features. This function is to resize the image and flatten the array to one dimension vector. The default value of the new size is 32*32.

The following to Fig. 3. is to test the bin spatial method.

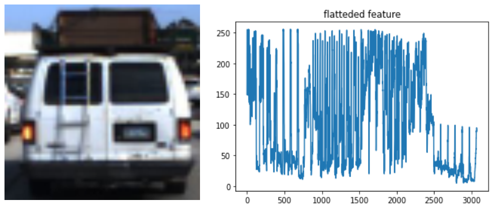

Fig. 3. Bin spatial output

Colour Histogram

When using the colour histogram, it is necessary to calculate the histogram of each channel. After finishing the statistics, using NumPy.concatenate() to connect three histograms. Here we need to pay more attention to the parameters bins and bins range in the function of the colour histogram. The value of the former is 32, and the range of the latter is (0,256). Fig. 4 is the output of testing the colour histogram function by using a simple training image.

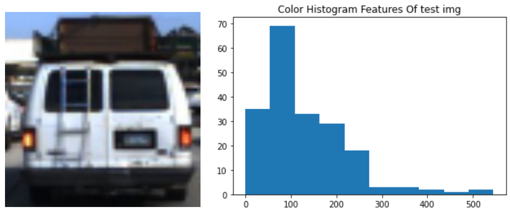

Fig. 4. Colour histogram output

Simple Image Feature

This function has one default value colour space which is RGB. The implementation of this function is to combine the previous three functions to get the whole features of a single image.

Extract Feature

This function has one default value colour space which is RGB. The implementation of this function is to use to for-loop combining with simple image feature to calculate

C. Training Classifier

Before training, using the sklearn shuffle to disorganize the order of training data. Then doing the feature extraction for all these features. Using the NumPy stack function to connect these two training datasets. Here one thing should be mentioned that using the StandardScaler function to standardize the data can increase the training efficiency and accuracy. Creating the label matrix with 0 or 1 is to label the output image in the classifier. Splitting the data as 8:2, which means 80% data for training and 20% data for testing. After finishing all these preparatory works, input the data into the linear SVM classifier and save the model as linear_SVM.mdl. Fig. 5. is the accuracy of this model, which is 97.69%.

Fig. 5. Accuracy of Linear SVM

D. Heat map and window sliding

In this sliding window section, three windows are set for window searching. The following Table. 1. will show the detail of these windows.

Table. 1. Window information

Window	Slide window	Overlap rate	y start	y end
Window1	(64,64)	0.75	301	501
Window2	(64,64)	0.5	311	571
Window3	(64,64)	0.25	321	591

Draw the heat map of the current test image 16th through the data as the Fig. 6. Showing below.

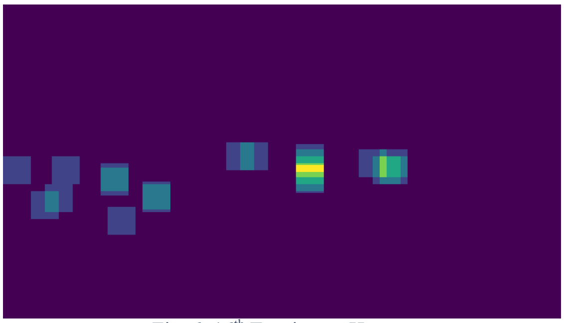

Fig. 6. 16th Test image Heat map

The corresponding accuracy rate and recall rate can be obtained by the proportion of the overlapping area. Here is the formula (2) for calculating the overlap area ratio.

$$recall\ rate = \frac{S_{overlap}}{S_{total}} = \frac{A \cap B}{A \cup B - A \cap B} \quad (2)$$

A, B in the above formula represents the area of two rectangular regions.

IV. Result And Discussion

A. Good Detection Example

Here is a good example of vehicle detection.

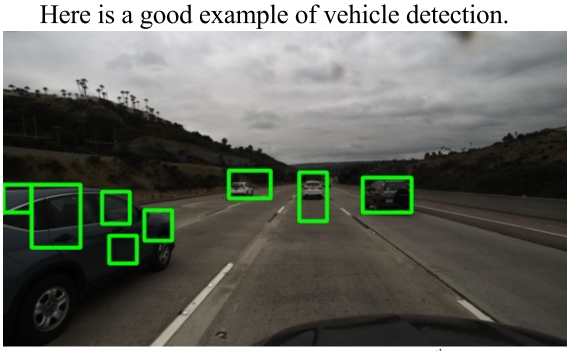

Fig. 7. Good Detection test image 16th

From Fig. 7. given above, the detection effect is pretty good. The three vehicles in front were all detected correctly. There are still two shortcomings. The first disadvantage is the uneven shape of the green box. This situation may be caused by the overlap rate and the size of the sliding window. The left car with several boxes. The car with multiple green boxes on the left may be caused by the following two situations. The first point is because the gradient is disturbed, the feature extraction is intermittent. The second point is that because the car does not completely appear in the picture. Therefore, the occlusion of the car causes the feature extraction to be inaccurate.

B. Poor Detection Example

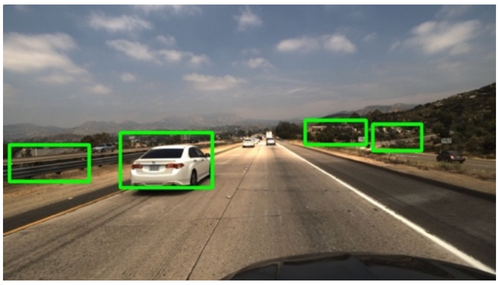

Fig. 8. Poor Detection Example

From Fig. 8. given above, three green rectangles do not select the vehicles but the dark part of the image. The HOG feature extraction is based on the gradient. This will cause the problem that the gradient is easily disturbed by noise information. It is not difficult to see from this picture that the noise in the surrounding background environment makes the HOG feature extraction get the wrong feature. Therefore, the background will be recognized as a vehicle.

C. Evaluation And Optimization

- Try to generate the non-vehicles without vehicles image in them. Reduce the impact of vehicle data in the non-vehicles ROI data collection
- Using AdaBoost to optimize the linear SVM [4]. The implement of AdaBoost is in the code file.
- Changing the window size, overlap rate, as well as the start and end location may increase the correctness of the detection.

V. Reference

[1] X. Wen, V. Bruno, S. Konrad, and P. Nicolas, “Street-side vehicle detection, classification and change detection using mobile laser scanning data”, Université Paris-Est. France, ISPRS Journal of Photogrammetry and Remote Sensing, 2016, Volume 114, pp. 166-178.

[2] Z. Wei, G. Shengyu, Z. Ling, and L. Xin, “Histogram of Oriented Gradients Feature Extraction From Raw Bayer Pattern Images”, Natural Science Foundation of China. Shanghai, vol. 67, pp. 946-950, May 2020.

[3] C. Xianbin, W. Changxia, Y. Pingkun, and L. Xuelong, “Linear SVM Classification Using Boosting HOG Features For Vehicle Detection in Low-Altitude Airborne Videos”, University of Science and Technology of China. Hefei, IEEE Conference, 2011, pp. 2469-2472.

[4] L. XuChun, W. Lei, and S. Eric, “AdaBoost with SVM-based component classifiers”, Faculty of Engineering and Information Sciences, University of Wollongong - Papers: Part A. 602.

[5] R. Sidheswar, R. K. Arun, and M. Chandrabhanu, “Analysis of various image feature extraction methods against noisy image: SIFT, SURF, and HOG”, Jiangsu University, Coimbatore, Tamil Nadu. India, ICECCT Conference, 2017.