Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Research on Driver's Distracted Behavior Detection Method Based on Multiclass Classification and SVM

Qingzhi Bu ✉, Jun Qiu ✉, Hao Wu , Chao Hu

Abstract—**To reduce the occurrence of traffic accidents caused by distraction. a detection method based on histogram of oriented gradient (HOG) and support vector machine (SVM) is proposed for driver's distraction behavior in this paper. Interest region of driver was detected first from video image, also the image was enhanced, denoised and normalized. Then the histogram of oriented gradient is used to extract the feature of the image. Meanwhile, the cross-validation method is used to optimize parameters in SVM. Finally, the effectiveness of the method is verified by compared with classical SVM algorithm and Local Binary Pattern algorithm (LBP) based on SVM algorithms. The results show that, the proposed method can obtain better classification accuracy.**

Index Terms—**Driver distraction, histogram of oriented gradient, cross validation, support vector machine**

I. Introduction

The Ministry of Public Security of the People's Republic of China show that there were more than 31.72 million newly registered motor vehicles, and the number of motor vehicles has reached 327 million, of which 240 million cars and 369 million motorists[1]. Traffic accidents caused by dangerous behaviors such as fatigue driving, distracted driving, drunk driving and speeding have become the world's most harmful[2]. According to insurance companies in Pennsylvania, twenty-five percent of motor vehicle accidents are caused by distracted driving[3], of which sixty-two percent of the total number of killed people has occurred by distraction or inattention, followed twelve percent of deaths are caused by using mobile phone while driving.Therefore, it is urgent to remind the driver's distraction behavior by effectively detection.

The driver's distraction behavior has always been a matter of concern to people, and it has also attracted more and more research for institutions and scholars at home and abroad. Wang[4] proposed a semi-supervised clustering algorithm for the detection of phone calls during driving, and got a good recognition rate. Liang[5] calculated the distance from the eyeball to the line of sight, steering wheel angle, lane position and turning error of steering wheel. Combined with these parameters, distracted driving and Concentrated driving features are given to support vector machine classifier to train and classify. Liu[6] and others proposed a Laplace support vector machine algorithm to detect the eyes and head of the driver in the semi-supervised mode to evaluate the driver's driving behavior. Cheng et al[7]. presented the image detection of driver's distraction, taking the yaw angle, pitch angle of the head and the offset of the nostril center as feature vectors, and then established the support vector machine classification model. Omid[8] have designed a wearable data acquisition system, which collects the galvanic skin response (GSR) signals of drivers when making phone calls and sending messages under natural conditions, and using convolution neural network (CNN) to learn and identify in real time in 2D spectrum and MEL cepstrum space. Wathiq O[9] used feed forward neural network to detect eye position, head position and mouth position to confirm whether the driver was distracted. Bhakti Baheti[10] et al. used the deep convolution neural network to detect the driver's distraction behavior, improved the VGG-16 model and achieved good results. Although the deep learning algorithm has a good effect on the driver's distraction detection, it is carried out in the case of large number of pictures, and for small sample image set recognition, the traditional machine learning algorithm has a certain advantage in training and recognition speed under the premise of ensuring the recognition rate.

In the propose work HOG algorithm along with SVM classifer are used. the features of the driver's distraction behavior using HOG algorirhm, and cross-validation is used to optimize SVM parameters, also a multi-class classifier is constructed for detection. The effectiveness of the proposed algorithm is verified by comparing with classical SVM algorithm and Local Binary Pattern algorithm (LBP) based on SVM algorithms.

II. THE PROPOSED METHOD

The driver distraction behavior image set is collected by State Farm company of the United States. the images provided contain many informations, so as to need to preprocess. then

This work is supported by Science and Technology Innovation Team Contruction Project of Ningbo in China under Grant 2014B82015, and general project of Education Department of Zhejiang in China under Grant Y201738805.

Qingzhi Bu is with the School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, 341000, China(1424052558@qq.com).

Jun Qiu is with the School of Information Science and Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo, 315100, China(qiujunnb@zju.edu.cn).

Hao Wu is with the School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, 341000, China(1424258177@qq.com).

Chao Hu is with the School of Information Science and Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo, 315100, China(huchao@nit.zju.edu.cn).

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

the HOG algorithm is used to extract the features, and the obtained feature vectors are input into the SVM for training. using the optimized training model classifying test set and the final test results are obtained. The overall structure of the system is shown in figure 1.

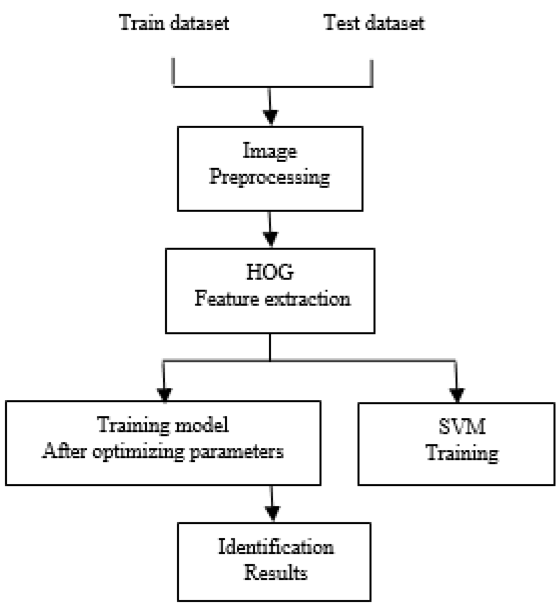

Fig. 1. General frame structure of the system.

A. Image preprocessing

Interest region in the image is acquired, and then the image is selected for graying, enhancement and denoising, in order to not only weaken the impact of external factors such as noise, light, shooting angle, but also enhance the effective information, the image set needs to be normalized to reduce the influence of some geometric changes and speed up the extraction.

B. Feature extraction

The HOG algorithm was first proposed at the CVPR conference by Dalal[11] in 2005. It was used to extract pedestrian features and achieved good results. The core idea of this algorithm is to calculate the gradient size and direction of a block in the image, accumulate the direction gradient histogram of the block, stack each block in the image, and generate the direction gradient histogram of the whole image, so that the feature of entire image can be described. The HOG extraction process is shown in figure 2.

a) Gamma normalization: In order to reduce the influence of illumination when the window slides, the image is grayed, then processed in Gamma space to reduce the sensitivity to illumination and shadowing.

b) Image gradient calculation: The calculation of image gradient not only can obtain some texture information, but also weaken the impact of illumination.

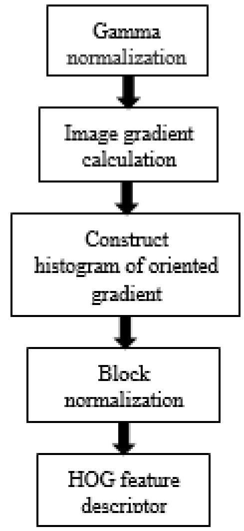

Fig. 2. HOG feature extraction process.

c) Construct histogram of oriented gradient: In practice this is implement by dividing the image window into small spatial region ("cells"), the cell has same size, the gradient direction is divided into nine directions, and the interval of each direction is $0° \sim 20°$. The gradient size of a pixel indicates the weight of the point. When the gradient direction of a pixel is in one of the nine directions, the weight of the pixel will be added to the interval of the direction.

d) Block normalization: The adjacent four cells form a block, and a block area contains gradient information. Due to the difference of exposure in some areas of the image, the contrast will be different, it will be resulting in a large range of gradient changes in the block area.Therefore, the block area needs to be normalized to compress the light and shadow information. The relationship between Block and cell is shown in Figure 3.

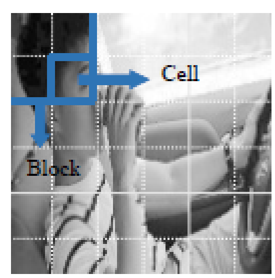

Fig. 3. Diagram of Block and Cell.

e) HOG feature descriptor: The gradient histograms in all block regions are connected to form HOG feature descriptors. The normalized image size is 128×128 pixels, which is the sliding window size. the size of a cell is 16×16 pixels, and 2×2 cell units form a block area. The sliding step size is 8×8 pixels, that is the total number of blocks is 225, and a cell has 9 directions, so the image has a total of 8100 feature vectors, we will refer to normalize descriptor blocks as

Histogram of Oriented Grident descriptor. The feature vector is input to support vector machine for training and recognition.

C. Support vector machine

The vector machine networks was proposed by Cortes C and Vapnik V[12] in 1995. It has outstanding performance in solving data classification problems in small sample, nonlinear and high dimensional feature space. The core idea is to find a separating hyperplane that will correctly distinguish the data. Support vector machines can be divided into linear and nonlinear.

f) Linear support vector machine: The basic idea of linear classifier is to find a optimal hyperplane in the target area for separable data to maximize margin between the vectors of classes. Assuming that there are m samples $(x_1,y_1)(x_2,y_1),,(x_m,y_m)$, where is y_m either 1 or -1, x is the q-dimensional feature vector, the purpose is to separate 1 from -1, so as to maximize the closest point between hyperplanes and x_m.

Hyperplane can be defined as

$$w^T x + b = 0 \tag{1}$$

It will show, that the weights w for optimal hyperplane in the feature space and x is feature vector, where b is constant. The support vector function is defined as

$$w^T x + b = \pm 1 \tag{2}$$

The distance between the support vector and the hyperplane is givend by

$$\frac{y(w^T x + b)}{||w||} = \frac{1}{||w||} = \frac{1}{\sqrt{w^T w}} \tag{3}$$

The linear function will accordingly be of the form:

$$f(x) = \mathrm{sign}(w^* x + b^*) \tag{4}$$

Where w* and b* are optimal solution. The distance from a support vector to the hyperplane is $1/||w||$, so the distance between the two support vectors is $2/||w||$. An optimal hyperplane is here defined as the linear decision function with maximal margin between the vectors of the two classes, see figure 4.

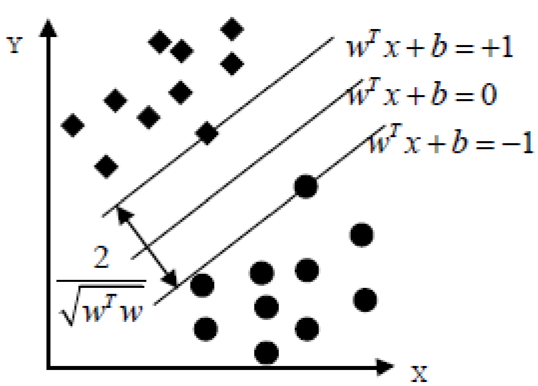

Fig. 4. Optimal hyperplane and support vector.

g) Nonlinear support vector machine: In the actual situation, nonlinear problem occur frequently, so using linear SVM classifier cannot completely distinguish the data set, which will cause a large problem of misclassification. Being sorely needed introduces a nonlinear support vector machine. The nonlinear classifier can display in figure 5.

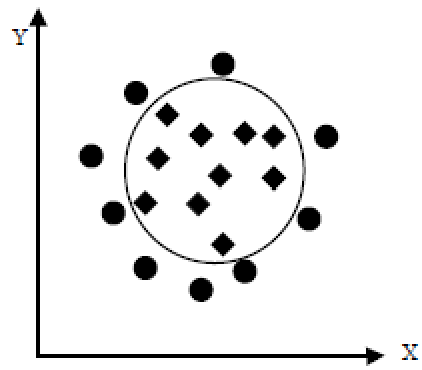

Fig. 5. Nonlinear classifier.

The idea of nonlinear classification is to transform the input from low dimensional space to high dimensional space through some mapping relation, and then to realize linear separability in high dimensional space[13]. The formula is as follows:

$$\begin{cases} \min\limits_{w,\xi}(\frac{1}{2}||w||^2 + C\sum\limits_{i=1}^{m}\xi_i) \\ s.t: y^{(i)}[w^T x + b] \geq 1 - \xi_i, i = 1,2,...,m, \xi_i \geq 0 \end{cases} \tag{5}$$

Where C is punishment factor and ξ is slack variable. Through a series of calculations, the final classification decision function is obtained:

$$f(x) = \mathrm{sgn}(\sum_{i=1}^{m} \beta_i y^{(i)} K(x, x^{(i)}) + b) \tag{6}$$

Where β is Lagrange multiplier, K(x,$x^{(i)}$) is kernel function and b is constant. The parameter gamma value and punishment factor C of the kernel function are the key factors to determine whether the classification is accurate. Usually, the parameter value is chosen by personal experience, which can not achieve the effect of specific classification,therefore, we need to use K-folk cross-validation method[14] to optimize the parameters C and Gamma. The principle of the algorithm be defined as the original data is divided into K groups, each group is tested respectively, the rest of the K-1 groups are used as training sets,after test K times, and the average recognition rate of K times is used as the property index of the classifier.

III. EXPERIMENTAL RESULTS

A. Experimental environments

The database from the Kaggle website[15]. Four categories of 1320 images were randomly selected, of which 1200 were used as training sets and the last 120 ones for test sets, including normal driving, texting, talking on the phone and drinking. The size of the original picture is 640 $\times$ 480 pixels, and it becomes 128 $\times$ 128 pixels after normalization. Meanwhile, we use QT5.6 software and OpenCV3.1 as the development environment in a computer, that is contains Inter Core TM i5-4200U CPU @ 2.3GHz and 8 GB RAM.

B. Comparison and Analysis of Algorithms

SVM classifier is usually adopted to solve a binary classification problem, the OpenCV algorithm in SVM based on the Libsvm, it uses one-versus-one method to deal with multi-class classification problem, so four classes of samples need 6 classifiers.

Using the HOG algorithm extract the features, and the 8100-dimensional feature vectors is input into the SVM classifier for training and recognition. First of all, in order to obtain the optimal parameters, the *C_SVC* classification model is established, and the radial basis kernel function (RBF) and the *K*-folk cross-validation method is adopted, where *K* is set to 10. To determine the value of parameter *C*, we set *C* from 0.0001 to 500 for searching for best value, which is searched in 5 steps each time. At the same time, the range of parameter Gamma is 0.0001 to 10, and the search is performed in 10 steps each time. the optimal value *C* is 39.062500 and the value of Gamma is 0.001500. Next, in order to verify the excellent performance of the proposed algorithm, The algorithm proposed in this paper is compared with classical SVM[16] and LBP based on SVM algorithm[17] under using the same database. The recognition results of various algorithms are shown in fig. 6 and table 1.

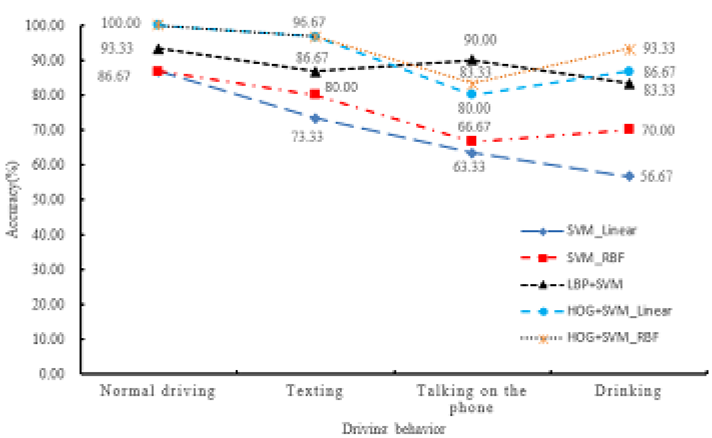

Fig. 6. HOG + SVM, SVM and LBP + SVM test sample result diagram.

TABLE I
RECOGNITION RESULTS OF DIFFERENT ALGORITHM MODELS

Method	*Average recognition rate (%)*
SVM_Linear	70.00[16]
SVM_RBF	75.84[16]
LBP + SVM	88.33[17]
HOG + SVM_Linear	90.84[18]
The proposed method	93.33

According to the data in table 1 and figure 6, it can be concluded that the SVM algorithm has the highest recognition rate of normal driving behavior in the four types of test samples, while the lowest recognition rate using linear SVM is drinking beverage behavior; LBP based on SVM algorithm has the highest recognition rate for normal driving behavior and the lowest recognition rate for drinking behavior; The HOG+SVM algorithm has the highest recognition rate for normal driving behavior and the lowest recognition rate for talking on the phone behavior. The average recognition rate of the model optimized by using HOG + SVM is higher than that of LBP + SVM algorithm and SVM algorithm. The above results show that it is feasible to use the HOG+SVM and optimize the parameters to detect the driver's decentralized behavior under certain error conditions.

IV. CONCLUSIONS

In this paper, a recognition algorithm based on HOG and SVM is proposed, which uses HOG to extract different features of driver's distracted behavior, and SVM classifier to classify decentralized behavior. Compared with SVM algorithm and LBP based on SVM algorithm respectively in the recognition rate of driver's distracted behavior. The experimental results show that using HOG+SVM and optimizing SVM parameters has a good effect on the recognition rate of this dataset, with an average recognition rate of 93.33%. During the experiment, the above results suggests that achieves the better detection accuracy in using the parameter optimization algorithm, However, the training and recognition speed still need to be improved, and more samples need to be used for training. It's will be seeking for a better algorithm to enhance the speed of training and recognition in the future, and the algorithm will be embedded into the Raspberry Pi with sound alerts, so as to realize deeper research on Driver's distracted behavior detection.

ACKNOWLEDGMENT

We would like to thank to the anonymous reviewers for their valuable comments and suggestions. This work is supported by Science and Technology Innovation Team Contruction Project of Ningbo in China under Grant 2014B82015, and general project of Education Department of Zhejiang in China under Grant Y201738805.

REFERENCES

[1] Lingfeng Jiang. the number of cars in the country exceeded 200 million for the first time in 2018 [online]. Available: http://www.mps.gov.cn/n2254098/n4904352/c6354939/content.html.

[2] Lei Zhao, "Research on driver fatigue detection method based on deep Learning and facial multi-source dynamic behavior fusion," Jinan, China:Shandong University, 2018.

[3] SafeStart. Top 10 causes distracted driving and what they all have common [online]. Available:https://safestart.com/news/top-10-causes-distracted-driving-and-what-they-all-have-common/.

[4] Jinru Wang, "Driver's call behavior detection based on semi-supervised support vector machine," Changsha, China: Hunan University, 2018

[5] Yulan Liang, Michelle L. Reyes, and John D. Lee, "Real-Time detection of driver cognitive distraction using support vector machines," IEEE Tansactions On Intelligent Transportation Systems, vol. 8, no. 2, pp.340-350, June 2007.

[6] Tian-chi Liu, Yan Yang, Guang bin Huang, et al, ”Driver distraction detection using semi-supervised machine learning,” IEEE Intelligent Transportation Systems Magazine, vol. 17,no. 4, pp. 1108-1120, April 2016.

[7] Wendong Chen, Rui Fu, Wei Yuan.et al, ”Image detection and grading warning of Driver’s Distraction,” Journal of Computer Aided Design and Computer Graphics. China, vol. 18, no. 8, pp. 1287-1296, 2016.

[8] Omid D, Mojtaba T, ”Driver distraction detection using MEL cepstrum representation of galvanic skin responses and convolutional neural networks,” 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, pp. 1481-1486, 20-24 Aug 2018.

[9] Wathiq O, Ambudkar B D, ”Driver safety approach using efficient image processing algorithms for driver distraction detection and alerting,” Intelligent Engineering Informatics.Advances in Intelligent Systems and Computing, vol. 695, pp. 461-469, 11 April 2018

[10] Bhakti Baheti, Suhas Gajre,Sanjay Talbar, ”Detection of Distracted Driver using Convolutional Neural Network,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1145-1151, 2018.

[11] Dalal N, Triggs B, ”Histograms of oriented gradients for human detection,”international Conference on computer vision and Pattern Recognition (CVPR). IEEE Computer Society, vol. 1, pp. 886-893, 2005.

[12] Cortes C, Vapnik V, ”Support-vector networks,” Machine learning, vol. 20, no. 3, pp. 273-297, 1995.

[13] Husheng Guo, Wenjian Wang, ”Granular support vector machine: a review,” Artificial Intelligence Review, vol. 51, no. 1, pp. 19-32, January 2019.

[14] Jayanthy S, Ayswaria S V, Vaishali M, et al, ”A HIGH QUALITY EMBEDDED SYSTEM FOR ASSESSING FOOD QUALITY USING HISTOGRAM OF ORIENTED GRADIENTS,”ICTACT Journal on Soft Computing, vol. 9, no. 1, pp. 1781-1787. January2018.

[15] StateFarm.State Farm Distracted Driver Detection.[online]. Available: https://www.kaggle.com/c/state-farm-distracted-driver-detection.

[16] Zhenhua Wang, Yinshan Jia, Xing Chen, ”Yinshan Jia. ”Research on driver fatigue detection based on SVM,” Science Technology and Engineering, China, vol. 11, no. 8, pp. 1828-1832, 2011.

[17] Likun Zhao, Xinhua Jiang, Haiyan Yang. ”Driver fatigue detection based on MLBPH-FF and SVM,” Computer and Modernization, China, no. 9, pp. 82-85 + 90, 2013.

[18] Zhengyou Wang, Xufeng Guo, Zewen Zhang, et al, ”Railway driver target detection based on local features,” China Sciencepaper, China, vol. 13, no. 20, pp. 2357-2364, 2018.