



A DRIVER DROWSINESS AND DISTRACTION WARNING SYSTEM BASED ON RASPBERRY PI 3 KIT

Dao Thanh Toan¹, Thien Linh Vo^{2,*}

¹University of Transport and Communications, No. 3, Cau Giay Street, Lang Thuong Ward, Dong Da District, Hanoi, Vietnam.

²University of Transport and Communications – Campus in Ho Chi Minh City, 450 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.

ARTICLE INFO

TYPE: Research Article

Received: 29/5/2019

Revised: 24/6/2019

Accepted: 23/8/2019

Published online: 15/11/2019

<https://doi.org/10.25073/tcsj.70.3.26>

* Corresponding author

Email: vtlinh@utc2.edu.vn

Abstract. In this article, a system to detect driver drowsiness and distraction based on image sensing technique is created. With a camera used to observe the face of driver, the image processing system embedded in the Raspberry Pi 3 Kit will generate a warning sound when the driver shows drowsiness based on the eye-closed state or a yawn. To detect the closed eye state, we use the ratio of the distance between the eyelids and the ratio of the distance between the upper lip and the lower lip when yawning. A trained data set to extract 68 facial features and “frontal face detectors” in Dlib are utilized to determine the eyes and mouth positions needed to carry out identification. Experimental data from the tests of the system on Vietnamese volunteers in our University laboratory show that the system can detect at real-time the common driver states of “Normal”, “Close eyes”, “Yawn” or “Distraction”.

Keywords: drowsiness detection, image sensing, HOG, Raspberry Pi 3, safety driver for Vietnamese.

© 2019 University of Transport and Communications

1. INTRODUCTION

Road traffic accidents have emerged as an important public health issue that needs to be tackled by multiple approaches. This incurs huge expenses for healthcare and causes frequent

overload conditions in the hospital system. The Global status report on road safety 2018, launched by WHO in December 2018, highlights that the number of annual road traffic deaths has reached 1.35 million. Road traffic injuries are now the leading killer of people aged 5-29 years. This number is expected to increase to about 1.9 million by 2030 and become one of the seven leading causes of human death [1].

Notice of March 28 of Vietnam National Traffic Safety Committee said in the first quarter of 2019, from December 16, 2018, to March 15, 2019, there were 4,030 traffic accidents, killing 1,905 people and injuring 3,141 people [2].

There are several causes of traffic accidents such as the operator's consciousness, use of the phone, drug or alcohol use, fatigue which leads to drowsiness and loss of concentration. So far, several techniques for detecting drowsiness have been studied [3-16], which can be divided into three basic directions as follows:

- The vehicle movement: abnormal change of vehicle in motion
- The physiology of the driver: using electroencephalogram (EEG) and electrocardiogram (ECG) signals.
- The behaviour of driver: external manifestations, facial expressions.

The first approach is quite complicated and costly since the vehicle needs to be equipped the sensing system to detect an abnormal change of vehicle in motion due to driver. On the other hand, even the second method brings about the most accurate result it requires a placement of sensed devices on the head, hands or chest that make the driver uncomfortable during driving [3]. The third approach is based on a non-touch measurement which has been considered to be an effective way to build a drowsiness detector. In addition, this method doesn't affect the driver due to its remote-sensing ability. In this paper, the drowsiness detection system focusing on the behaviour of driver is presented. The drowsiness detection system developed in our University laboratory consists of three components: the Raspberry Pi 3 Kit, Pi Camera module and a speaker to emit warning sounds. A camera is mounted in the vehicle to capture the driver's face and constantly monitor the driver's eyes and mouth. The Raspberry Pi 3 kit is responsible for analysing the frames continuously and warning the driver in real time if there is an abnormal detection so that the driver can focus again [6,7]. Thanks to its small size, it can be easily equipped in any type of vehicle. In addition, this system is cheaper than other safety measures that are equipped in vehicles or equipped with drivers.

2. METHODS OF FACE IDENTIFICATION

2.1. Haar cascade classifier

The Viola-Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones [8]. Although it can be trained to detect a variety of object classes, it was motivated primarily by the problem of face detection. The training method is basically

matching all available Haar-like features into a grayed and standardized original image (resizing as required). The appropriate Haar-like features will be extracted and selected the most optimal threshold according to the 8-bit gray level. The advantage of using features instead of raw pixel values is the possibility to compensate little variations in the appearance of the object which makes classification easier. The computation of these features is based on the comparison of pixel intensities [8].

All human faces share some similar properties. These regularities may be matched using Haar Features. A few properties common to human faces:

- The eye region is darker than the upper-cheeks.
- The nose bridge region is brighter than the eyes.
- Composition of properties forming matchable facial features:
- Location and size: eyes, mouth, bridge of the nose.
- Value: oriented gradients of pixel intensities.

The four features matched by this algorithm are then sought in the image of a face as shown in Fig. 1. All trained Haar-like features will be scanned over the input image, i.e. every pixel in the image will at least once slide through by Haar-like, the matched areas which result in the same with many Haar-like features will be marked and identified as the face.

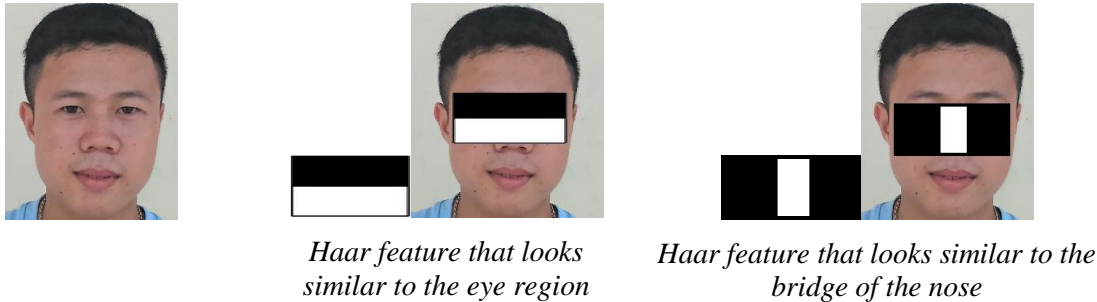


Figure 1. Matching Haar features.

2.2. Histogram of Oriented Gradients (HOG)

A tool of HOG describing functional features creates other forms of objects in vector space by extracting the HOG features (HOG descriptors) of that object. It restricts the information that is not useful or highlighted the object border by the intensity gradient feature of the object boundary [10]. For the human face recognition problem, this useful information will go through the SVM (Support Vector Machine) classifier, the output will predict the result in the image with the face or not. Therefore, HOG is mainly used to describe the shape and appearance of an object in the image. The essence of the HOG method is to use information about the distribution of gradient intensities or Edge Directions to describe objects in the image. The HOG operators are implemented by dividing an image into cells, each cell (8×8) will draw a histogram of the oriented gradients for points within the cell.



Figure 2. Sample face image (left) and image of HOG descriptor (right).

To enhance the recognition performance, the histograms will normalize the contrast by calculating an intensity threshold in a region larger than the cell, called blocks (4 cells) and using that threshold value to standardize all cells in the block. Results after the normalization step help HOG capable of high stability with variation in brightness across the image as presented in Fig. 2.

2.3. Proposed Method

Haar method gives a faster speed of the object detection, but it is affected by the fluctuated light intensity. That limitation can be overcome by using HOG classifier since the HOG works on the principle of segmentation [17]. The HOG determines the face of driver with respect to a tilted or distorted state, allowing a detection of the human face at high accuracy. Therefore, we employ the HOG classifier running on Raspberry Pi 3 to develop in the system.

3. SYSTEM CREATION

For implementation, we use Python language in collaboration with the open source libraries OpenCV and Dlib on the hardware platform of Raspberry Pi 3 model B+. The system to identify driver drowsiness and distraction state operates according to the following steps:

- Record of video source from the camera in the cockpit;
- Identification of the driver faces from the received videos;
- Localization of the eyes and mouth spots on the face;
- Calculation of the ratio of the eye and mouth and comparison the ratio value with the predetermined threshold value (the ratio of eye ratio within the closed eye area);
- Detection or prediction of the drowsiness and distraction;
- Generation of a warning sound if drowsiness and distraction are detected.

In short, we scanned the face with the HOG tool combined with the SVM classifier to determine whether there is a human face or not. Once a face is identified, we continued to identify facial points by using the Dlib library as shown in Fig. 3. The Dlib library supports 68 facial recognition points implemented through the Facial Landmark function [8,14]. To locate such 68 points on the human face, Dlib's Facial Landmark set was trained with the iBUG 300-W dataset set, the input training data set is 1000 human face images with the 68 points marked manually [8].

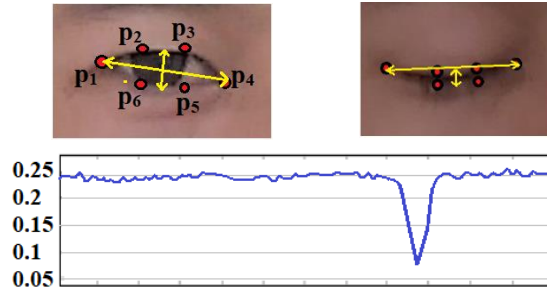


Figure 3. Distance ratio of open and closed eyes taken in our laboratory.

Next, the system will calculate the ratio of the distance between two eyelids (E) according to the below formula [10,14]:

$$E = \frac{|P_2 - P_6| + |P_3 - P_5|}{2|P_1 - P_4|} \quad (1)$$

Where the P_1, P_2, P_3, P_4, P_5 and P_6 represented the eye factors of the driver as illustrated in Fig. 4. Following to equation (1), the E value will be reduced whenever the eye closes (see Fig. 4). From experiments on the 100 Vietnamese people, we found that the threshold of eye-closing is about 0.22. The system then was trained to activate the warning functionality once the E below a certain threshold of 0.22.

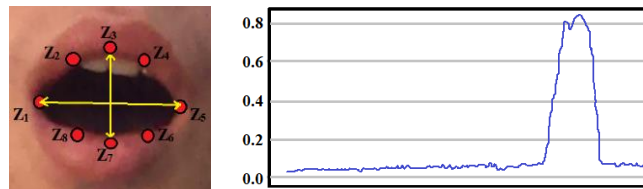


Figure 4. Distance ratio of open and closed mouths taken in our laboratory.

On the other hand, the system will also determine the ratio of distance (M) based on the state of open or closed mouth shown in Fig. 4 using following equation [10,13,16]:

$$M = \frac{|Z_2 - Z_8| + |Z_3 - Z_7| + |Z_4 - Z_6|}{3|Z_1 - Z_5|} \quad (2)$$

Where the $Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7,$ and Z_8 represented the mouth factors of the driver as illustrated in Fig. 4. Similar to the E , the M value goes down when the mouth is in close. The M was found to be about 0.27 from experiments on the 100 Vietnamese people.

Fig. 5 and Fig. 6 show the algorithm flowchart and the system model based on the hardware platform of Raspberry Pi 3, respectively. The system captures a frame from the camera and determines the location of the face in the frame. Next, the system identifies the basic points on the face to find the eyes and mouth. These points are calculated to identify and to compare with threshold values. The system synthesizes and gives results on whether the driver is drowsy or distracted. The threshold values are selected and adjusted from Vietnamese drivers.

4. EXPERIMENTAL RESULTS

Fig. 7 presents the experimental results for a normal case from a 4-Vietnamese-driver with a 4-test-state. A “normal case” here is meant the driver is without class, hat, or any other belongs on the face. As can be seen, there are 3 states showing a warning signal for the situation of “Close eyes”, “Yawn”, and “Distraction” [15,16]. In common case, the total of experimental frames is 200 in both well-lighted and poorly-lighted environments. Warning signals will be given for three cases: Closed eyes, Yawn, Distraction as summarized in Table 1. Overall, at a well-light environment, the system can be recognized completely exactly the state of the car driver (100 %). And even when the lighting condition at poor, the accuracy of the system detector is as high as more than 86 %.

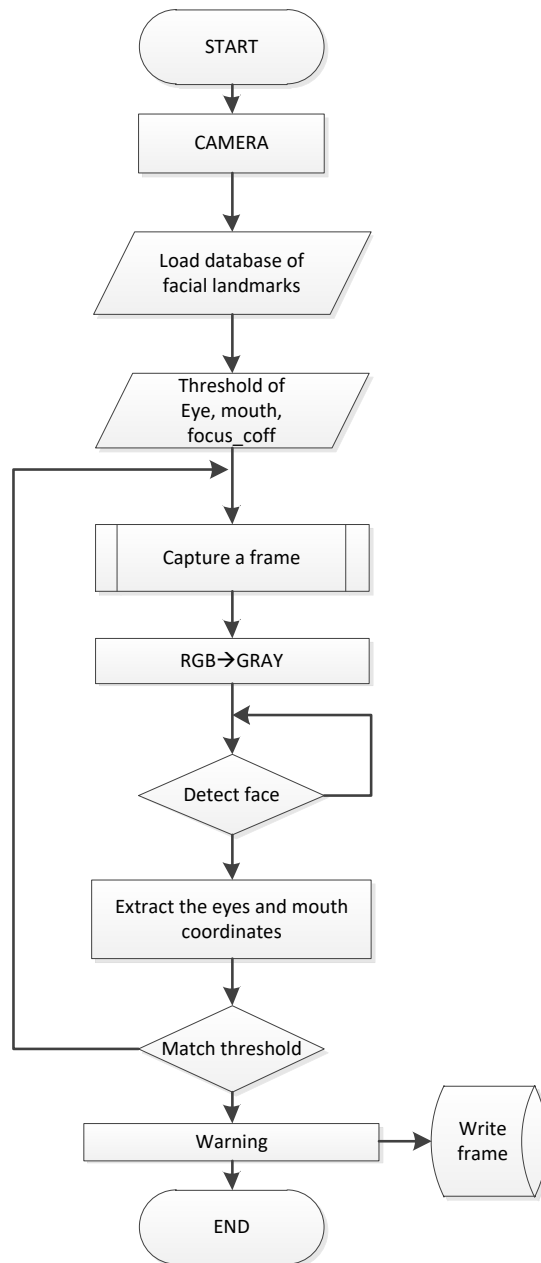


Figure 5. Algorithm flowchart of detector system.



Figure 6. Experimental system model fabricated in our University laboratory.

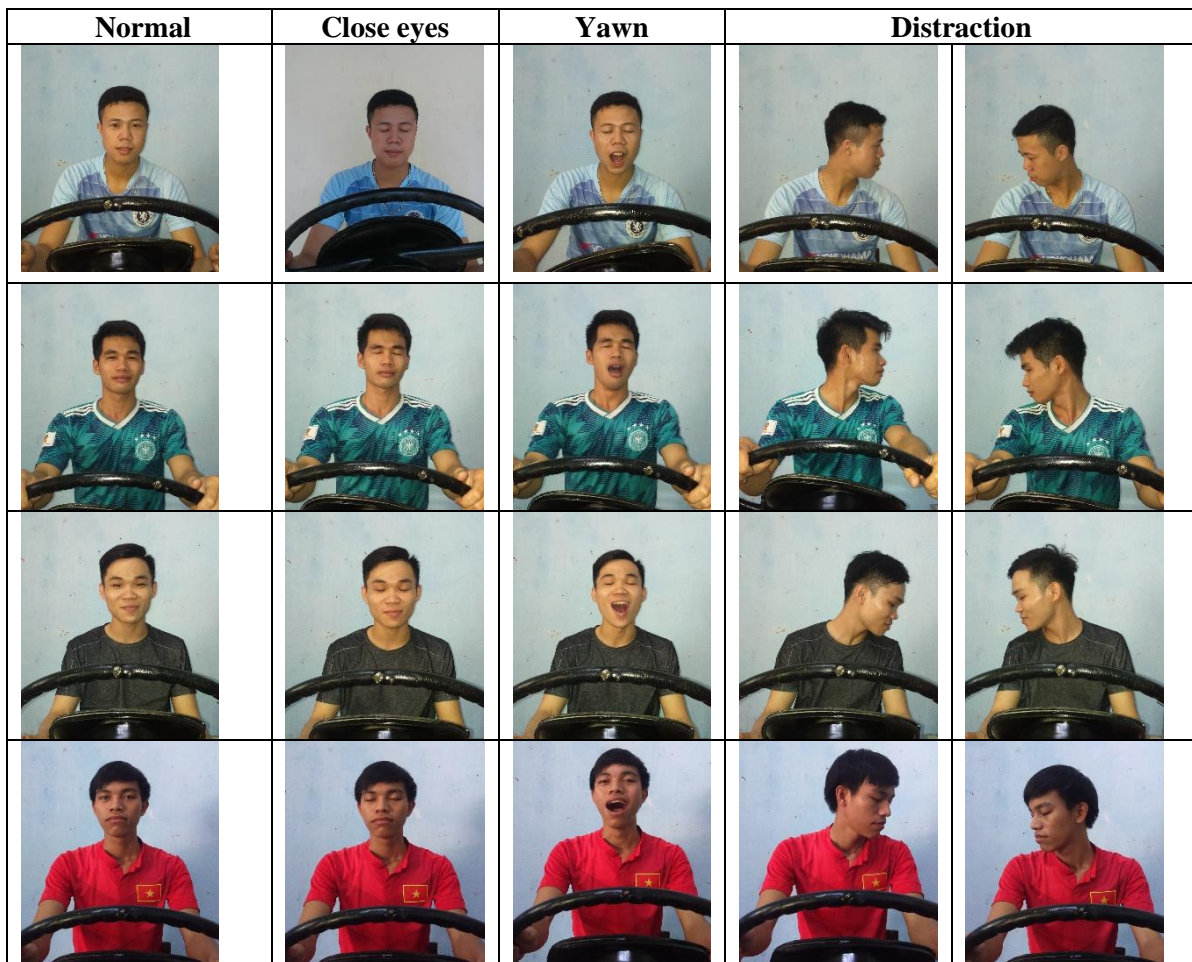


Figure 7. Photos extracted from experimental frames in normal case.

Table 1. Experimental results from normal case.

Lighting conditions	Number of frames	Close eyes		Yawn		Distraction	
		Number of correct identification frames	Accuracy	Number of correct identification frames	Accuracy	Number of correct identification frames	Accuracy
well-lighted	200	200	100%	200	100%	200	100%
poorly-lighted	200	172	86%	200	100%	190	95%

On the other hand, when the car driver at a “special case” such as wearing a glass or mask as shown in Fig. 8. Using the same process, the total of experimental frames is 40 frames and the system still can detect and active the warning signal for the states of “Closed eyes”, “Yawn”, and “Distraction”. In this case, the accuracy value of the system detector is lower than that in the normal case as shown in Table 2. However, the system developed here still is able to give an warning signal when the car driver at such three abnormal states.



Figure 8. Photos extracted from experimental frames in special case.

Table 2. Experimental results from special case.

Lighting conditions	Number of frames	Close eyes		Yawn		Distraction	
		Number of correct identification frames	Accuracy	Number of correct identification frames	Accuracy	Number of correct identification frames	Accuracy
well-lighted	40	35	87%	28	70%	40	100%
poorly-lighted	40	32	80%	25	62,5%	190	95%

7. CONCLUSION

We have demonstrated a detector of driver drowsiness and distraction based on the Python language in combination with the open source libraries OpenCV and Dlib on the hardware platform of Raspberry Pi 3 model B+. The HOG classifier is used to calculate the ratio of eye and mouth points. The test system operates at a relatively high accuracy. In all cases of well-lighted or poorly-lighted environments with/without glasses or masks, the system is able to generate a warning sound when the driver shows drowsiness based on the eye-closed state or a yawn. The experimental study presented here can be in the first step to domestically fabricate a low-cost drowsiness and distraction detector system for Vietnamese which will be equipped in a car, resulting in a contribution in reduction of traffic accident on road.

ACKNOWLEDGMENT

We would like to thank the students of Faculty of Electrical-Electronic Engineering, University of Transport and Communications (HCMC campus) for their helps in performing testing.

REFERENCES

- [1] WHO. Road Safety. The global status report on road safety 2018. https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/
- [2] Vietnam National Traffic Safety Committee, 2019. <http://tuyengiao.vn/uy-ban-an-toan-giao-thong/trong-quy-i2019-xay-ra-tren-4-000-vu-tai-nan-giao-thong-120125>. (In Vietnamese)
- [3] A. Sahayadhas, K. Sundaraj, M. Murugappan, Detecting driver drowsiness based on sensors: A review, *Sensors*, 12 (2012) 16937-16953. <https://doi.org/10.3390/s121216937>
- [4] M.A, Assari, M. Rahmati, Driver drowsiness detection using face expression recognition, in *Proceedings of the IEEE International Conference on Signal and Image Processing Applications*, Kuala Lumpur, Malaysia, 337–341, 2011. DOI: 10.1109/ICSIPA.2011.6144162
- [5] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, S.C. Jun, Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data, *Front. Hum. Neurosci.*, 10 (2016) 219. <https://doi.org/10.3389/fnhum.2016.00219>
- [6] N. Agrawal, S. Singhal, Smart drip irrigation system using raspberry Pi and Arduino, in *Proceedings of International Conference on Computing, Communication and Automation*, Noida, India, 928-932, 2015. DOI: [10.1109/CCAA.2015.7148526](https://doi.org/10.1109/CCAA.2015.7148526)
- [7] J. Marot, S. Bourennane, Raspberry Pi for image processing education, in *Proceedings of 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 2017. DOI: [10.23919/EUSIPCO.2017.8081633](https://doi.org/10.23919/EUSIPCO.2017.8081633)
- [8] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, 2001. DOI: [10.1109/CVPR.2001.990517](https://doi.org/10.1109/CVPR.2001.990517)
- [9] V. Kazemi, J. Sullivan, One Millisecond Face Alignment with an Ensemble of Regression Trees paper, in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Washington DC, USA, 1867-1874, 2014, DOI: [10.1109/CVPR.2014.241](https://doi.org/10.1109/CVPR.2014.241)
- [10] N. Dalal, B. Triggs, Histogram of Oriented Gradients for Human Detection, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, 2005. DOI: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)
- [11] C. Meng, L. Shi-wu, S. Wen-cai, G. Meng-zhu, H. Meng-yuan, Drowsiness monitoring based on steering wheel status, *Transportation Research Part D: Transport and Environment*, 66 (2019) 95-103. <https://doi.org/10.1016/j.trd.2018.07.007>
- [12] S.Arefnezhad, S. Samiee, A. Eichberger, A. Nahvi, Driver Drowsiness Detection Based on Steering Wheel Data Applying Adaptive Neuro-Fuzzy Feature Selection, *Sensors* 19 (2019) 943.
- [13] G. Li, C. Wan-Young, Combined EEG-Gyroscope-tDCS Brain Machine Interface System for Early Management of Driver Drowsiness, *IEEE Transactions on Human-Machine Systems*, 48 (2018) 50-62. DOI: [10.1109/THMS.2017.2759808](https://doi.org/10.1109/THMS.2017.2759808)
- [14] B.-G. Lee, B.-L. Lee, W.-Y. Chung, Wristband-type driver vigilance monitoring system using smartwatch, *IEEE Sensors Journal*, 15 (2015) 5624–5633.
- [15] M. R. Guedira, A. El Qadi, M. R. Lrit, M. E. Hassouni, A novel method for image categorization based on histogram oriented gradient and support vector machine, in *Proceedings of the International Conference on Electrical and Information Technologies*, Rabat, Morocco, 2017. DOI: [10.1109/EITech.2017.8255229](https://doi.org/10.1109/EITech.2017.8255229)
- [16] T. Soukupova, J. Cechin, Real-Time Eye Blink Detection using Facial Landmarks, in *Proceedings of the 21st Computer Vision Winter Workshop*, Rimske Toplice, Slovenia, 2016.