A Fatigue Driving Detection Method based on Deep Learning and Image Processing

Zhong Wang$^{1,2}$, Peibei Shi$^{1,4}$ and Chao Wu$^3$

$^1$School of Computer Science and Technology, Hefei Normal University, No. 1688 Lianhua Road, Hefei, Anhui, China.
$^2$Anhui Province Key Laboratory of Big Data Analysis and Application, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China.
$^3$Network and Information Center, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei, Anhui, China

Email: zhongw@ustc.edu.cn; *Corresponding Author: pbshi@hfnu.edu.cn;

Abstract. Driving fatigue is one of the important causes of traffic accidents. It is of great significance to study fatigue driving detection algorithms to improve human life and property safety. This paper proposes a fatigue driving detection method based on deep learning and image processing. First, the driver's face image is obtained in real time through the camera and the face image is detected using the MTCNN model. Next the image processing is performed on the face image, including three steps: grayscale processing, binarization processing, and human eye detection. Then we check the legitimacy of the human eye image and calculate the eye closure rate, and finally use the PERCLOS principle to analyze the fatigue state of the driver. The experimental results show that the proposed method has high detection rate and low false alarm rate, and has strong practicality.

1. Introduction

Driver fatigue driving is one of the important causes of traffic accidents, mainly manifested in drowsiness, closed eyelids, decreased attention and other symptoms. In the past two decades, the transportation departments of countries all over the world have invested more manpower and material resources in fatigue driving detection, and have achieved some effects, but they have not yet achieved good applications [1]. Recently, many car companies such as Volkswagen, Toyota, and Nissan have also installed driver assistance modules in their vehicles to prevent traffic accidents. According to a study by the U.S. Highway Safety Administration, 22% to 24% of traffic accidents are related to the mental state of the driver. In China, traffic accidents caused by fatigue driving account for 20% of the total number of accidents each year, and account for more than 40% of major traffic accidents. A Canadian study reported that 20% of fatal crashes are related to fatigue [2]. Records show that 34% of road accidents in Pakistan are related to fatigue driving [3]. There are also reports that 20% of traffic accidents in the European Union are due to fatigue driving [4]. Therefore, research on fatigue driving detection algorithms is of great significance to reduce traffic accidents and protect human life and property safety.

Fatigue driving detection technology is mainly divided into three categories: detection methods based on physiological characteristics, detection methods based on vehicle characteristics, and detection methods based on visual characteristics. Physiological characteristics-based detection methods mainly collect driver's physiological signals to detect driver fatigue, such as electroencephalography (EEG) [5], electrocardiogram (ECG) [6], electro-oculography (EoG) [7] and
surface electromyogram (sEMG). This method is contact detection, which is very intrusive, so it is limited in real-time applications. Detection methods based on vehicle characteristics can be divided into three types of methods based on steering wheel angle, lane deviation, and posture change. This method is non-invasive, but the detection result is greatly affected by the driving habits of the driver. The visual-based feature detection method is to realize the fatigue detection of the driver by analyzing the physical features of the driver's eye state, mouth movement and head posture. Different from the previous two methods, this method has the advantages of non-contact, non-interference, accuracy and so on.

Detection methods based on visual features are further divided into three categories based on their physical characteristics: 1) Detection methods based on eye features. In this method, the closed eyes rate, eyelid distance, and open eyes percentage were used as evaluation indexes. Sigari et al. [8] uses Haar-like feature detector to detect human faces, uses template matching method to extract features from images, and tests in the actual driving environment. The accuracy of this method is greatly affected by the influence of light. Mandal et al. [9] proposed a fatigue detection system based on bus drivers, which detects fatigue status by estimating eye opening and PERCLOS. This method uses HOG features to extract head and shoulder features, and uses SVM classifiers for fatigue detection. 2) Detection method based on mouth features. Yawning and mouth opening can also be good indicators of fatigue detection. Alioua et al. [10] implemented yawning to detect driver fatigue, and used a yawning counter to analyze driver fatigue. Jie et al. [11] proposed a method for detecting yawns based on geometric and appearance features of the mouth and eye regions. 3) Detection method based on head posture. When a driver is fatigued, the number of dozing and nodding will increase significantly. Ruiz et al. [12] proposed a method to determine the head posture by training a convolutional neural network, and used RGB combined with classification and regression loss to predict the angular change of the head posture.

Traditional visual feature-based detection methods mainly use Haar-like features, HOG features, SIFT features, etc. as feature descriptors, and classifiers mainly use SVM classifiers, integrated classifiers, and so on. With the development of computer vision, deep learning has made great progress in the field of object detection. CNN-based fatigue detection has become a current hotspot. For example, Zhang et al. [13] uses facial key points to extract eye images, and uses CNN to determine eye status to detect fatigue driving. This paper combines the advantages of deep learning and image processing methods. First, face detection is performed using deep learning methods, and then eye positioning and analysis are performed using image processing methods.

2. Methods

Figure 1 gives the overall framework of the proposed method. First, the driver's face image is obtained through the camera, and the face image is detected using the MTCNN [14] model. Next the image processing is performed on the face image, including three steps: grayscale processing, binarization processing, and human eye detection. Then we check the legitimacy of the human eye image and calculate the eye closure rate, and finally use the PERCLOS principle to analyze the fatigue state of the driver.

MTCNN

Camera

Image

Face

Image

grayscale

processing

binarization

processing

human eye
detection

Eye

Image

legal check

fatigue

analysis

Fatigue

State

Figure 1. The framework of the proposed method

2.1. Face Detection

Face detection is completed by MTCNN[14] model. This method is a cascade and coarse-to-fine deep learning model. It is also a very widely used detector, which can be deployed on traditional hardware platform, and has certain practical value for fatigue driving detection. The MTCNN model is a cascade
of three convolutional neural networks, while face detection, feature point positioning is also completed. The real model is divided into three stages. The first stage uses a shallow CNN network to quickly generate a series of candidate windows. The second stage uses a more capable CNN network to filter most non-face candidate windows. In the third stage, the 5 marked points on the face are found through a more capable network. The complete MTCNN model detection process is shown in Figure 2.

Figure 2. The detection process of MTCNN

2.2. Image Processing
Image processing includes three steps: grayscale processing, binarization processing and human eye detection.

2.2.1. Grayscale processing. Color image is transformed into grayscale image, and the formula of transforming RGB color space into grayscale image is:

\[
gray = 0.299 \times r + 0.587 \times g + 0.114 \times b
\]

We count the number of times 0 to 255 gray levels appear and get the gray histogram. Considering the driving situation of the vehicle, the general driver's left face has high brightness. Therefore, the following simple gray scale transformation is performed to calculate the left average gray scale and the right average gray scale. The parameters are adjusted to reduce the brightness of the left and right faces. The result is shown in Figure 3.

Figure 3. The flow diagram of grayscale processing

2.2.2. Binarization processing. First, use the gray histogram to calculate the gray level, set the threshold and binarize the image. The smaller than the threshold is black, and the others are white. As can be seen from the figure below, the hair is a large black, and it is basically not connected to the eye part, which can be removed by flood filling method. Finally, a vertical gradient projection is performed. The vertical gradient is the gray difference between the pixel and the pixel in its vertical position. The results are shown in Figure 4.
2.2.3. **Eye detection.** Eye detection includes vertical detection, horizontal detection, accurate detection and eyelid detection. The results are shown in Figure 5.

**Horizontal detection:** Differentiate the left and right eyes and determine the vertical position, first store the binary area of vertical detection in the array, find the maximum value max and its position imax, set the threshold threshold = max / 10, and then traverse the left and right eyes from the middle of the array, marking the first time a non-zero point is encountered and the second time a zero point is encountered, the difference between the two points is the left eye width or right eye width.

**Vertical detection:** Locate the approximate horizontal position of the eyes, and a pair of obvious peaks and valleys will appear where the eyes and eyebrows appear from the vertical gradient projection of the binary image.

**Accurate detection:** Save the data of the above steps to an array, starting from the bottom of the array, starting at the point where the value is greater than 0, and stopping at the point where the value is 0, so as to obtain the precise range of the human eye.

**Eyelid detection:** The precise range of the human eye can be regarded as a rectangular frame, and the maximum height value of the rectangular frame is not 0 as the eyelid opening degree.

2.3. **Fatigue Detection**

Fatigue detection is divided into three steps: data legality detection, data processing and data analysis.

**Data legitimacy check:** The human eye image is checked for legitimacy to remove illegal data. The rules are: if the left and right eyelids are open to a large extent or the absolute value of the straight line slope across the center of the eyes is greater than 0.5, the data is considered illegal.

**Data processing:** First calculate the current eyelid opening degree, take the average of the left and right eyelid opening degrees, and save them in the eyelid sequence, and then calculate the eyelid opening degree over a period of time, remove the largest 5% of them, and use the remaining maximum as the maximum binocular opening degree.

**Data analysis:** The length of time the eyes are closed is closely related to the degree of fatigue. The longer the driver's eyes are closed, the more severe the degree of fatigue. Therefore, the degree of driving fatigue can be determined by measuring the length of time the eyes are closed. The paper uses PERCLOS as the index of fatigue detection. PERCLOS refers to the percentage of closed time of eyes in a certain time, which is divided into three standard types: P70, P80, and EM. Compared with other methods, PERCLOS can more accurately and intuitively reflect the driver's fatigue state. This paper adopts the P80 standard, which is the indicator that is generally considered to reflect the fatigue program of a person. The calculation formula of PERCLOS is:
Where ECF represents the number of eyes-closed frames and AF represents the total frame number.

3. Experimental Results
In order to verify the effectiveness of the algorithm, we self-built a fatigue detection data set and collected several test videos through the camera, including videos of 5 testers' fatigue and normal state under different lighting conditions. First we use MTCNN model to detect human faces, and then perform image processing. The data set of the MTCNN face detection model comes from multiple public face data sets. We select face images of different scales, poses, expressions, and illuminations to form a new data set. The experiment uses the deep learning framework TensorFlow to implement the network. The data set is divided into 2 subsets, and training and testing are performed according to 3: 1.

Table 1. Detection Results of the proposed method.

| ID | Closed frames detected | Actual closed frames | Accuracy | PERCLOS value | Fatigue State |
|----|------------------------|----------------------|----------|---------------|---------------|
| 1  | 36                     | 35                   | 97%      | 0.15          | no            |
| 2  | 125                    | 120                  | 96%      | 0.45          | yes           |
| 3  | 52                     | 48                   | 92%      | 0.12          | no            |
| 4  | 87                     | 82                   | 94%      | 0.52          | yes           |
| 5  | 75                     | 70                   | 93%      | 0.36          | yes           |

Table 1 shows the accuracy of the human eye state detection and the fatigue state of the five testers. As can be seen from the table, the algorithm proposed in the paper has a high accuracy rate in real environments. The accuracy rate of each tester is more than 90%, and the average accuracy rate of the algorithm is close to 95%. In addition, Table 1 also gives the PERCLOS value of each tester. It can be seen that if the PERCLOS value is higher than 0.25, it is considered a fatigue state. Combining test videos with actual samples, the detection results given in the paper are consistent with the actual samples. Figure 6 shows the detection results of a single tester. The left side is the result of face detection and eye positioning, and the right side is the sequence diagram of the eye opening degree.

![Figure 6](image-url)

Figure 6. The results of fatigue state

4. Conclusions
This paper proposes a fatigue driving detection algorithm based on deep learning and image processing. Face detection using MTCNN model in deep learning. The image processing method includes three steps of grayscale processing, binarization, and human eye detection. The driver's fatigue status is estimated from the PERCLOS values of multiple videos. Experimental results verify that the proposed algorithm is informative.

5. Acknowledgments
This work was supported by the National Natural Science Foundation of China (61976198), the Natural Science Research Key Project for Colleges and University of Anhui Province (KJ2018A0498,
KJ2019A0726), and the Anhui Province Key Laboratory of Big Data Analysis and Application Open Project.

6. References

[1] Sikander G, Anwar S. Driver fatigue detection systems: A review [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20 (6): 2339-2352.

[2] (2011). Road Safety in Canada. Accessed: Mar. 24, 2017. [Online]. Available: https://www.tc.gc.ca/.

[3] Azam K, Shakoor A, Shah R A, et al. Comparison of fatigue related road traffic crashes on the national highways and motorways in Pakistan [J]. Journal of Engineering and Applied Sciences, 2014, 33(2):47-54.

[4] Fatigue. Accessed: Jan. 21, 2017. [Online]. Available: https://ec.europa.eu/transport/road_safety/.

[5] Chai R, Naik G R, Nguyen T N, et al. Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system [J]. IEEE journal of biomedical and health informatics, 2016, 21 (3): 715-724.

[6] Tsuchida A, Bhuiyan M S, Oguri K. Estimation of drowsiness level based on eyelid closure and heart rate variability [C]//2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2009: 2543-2546.

[7] Zhang Y F, Gao X Y, Zhu J Y, et al. A novel approach to driving fatigue detection using forehead EOG [C]//2015 7th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2015: 707-710.

[8] Sigari M H, Fathy M, Soryani M. A driver face monitoring system for fatigue and distraction detection [J]. International journal of vehicular technology, 2013:1-12.

[9] Mandal B, Li L, Wang G S, et al. Towards detection of bus driver fatigue based on robust visual analysis of eye state [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 18 (3): 545-557.

[10] Alioua N, Amine A, Rziza M. Driver’s fatigue detection based on yawning extraction [J]. International journal of vehicular technology, 2014.

[11] Jie Z, Mahmoud M, Stafford-Fraser Q, et al. Analysis of yawning behaviour in spontaneous expressions of drowsy drivers [C]//2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). IEEE, 2018: 571-576.

[12] Ruiz N, Chong E, Rehg J M. Fine-grained head pose estimation without keypoints [C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018: 2074-2083.

[13] Zhang F, Su J, Geng L, et al. Driver fatigue detection based on eye state recognition [C]//2017 International Conference on Machine Vision and Information Technology (CMVIT). IEEE, 2017: 105-110.

[14] Zhang K, Zhang Z, Li Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. IEEE Signal Processing Letters, 2016, 23 (10): 1499-1503.