ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

Research on Fatigue Driving Detection Under Wearing Mask

Shaosong Duan[1, 2, a], Lecai Cai[2, b, *] and Xing Liu[1, 2, c]

[1]School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China

[2]Sanjiang Research Institute of Artificial Intelligence and Robotics, Yibin University, Yibin, China

[a]1416508890@qq.com, [b]ybxyclc@163.com, [c]1610830030@qq.com

Abstract

The popularity of COVID-19 has rapidly affected our daily life. One of the solutions to prevention is to wear masks. At present, wearing masks has become a new norm. However, when a motor vehicle driver wears a mask, all the parts below his eyes are blocked. Most face detection algorithms can't extract effective information, which leads to low accuracy and completely unreliable precision, and can't achieve a fatigue detection effect. Therefore, in order to reduce the traffic accidents caused by this kind of fatigue driving, this paper first detects the face from the image, then accurately locates the pupil in the detected face area, compares some non-deep learning algorithm face detectors, using the Residual Network (ResNet) to connect the SSD network to detect the face key points. The front structure of the network uses the residual module to construct the deep learning face detection algorithm of ResNet10SSD. The pupil circle detection algorithm of Hough transform is adopted to analyze the pupil opening of the eyes, and the driver's eye opening and closing state is effectively judged. Finally, the blink frequency and PERCLOS value are obtained. The experimental results show that this method can effectively detect the eye indicators under the condition of face occlusion, and it has good universality and effectiveness, thus further realizing fatigue monitoring.

Keywords

Fatigue detection; Face detection; Eye features; PERCLOS.

1. INTRODUCTION

The number of motor vehicles in China has exceeded 200 million. The increase of vehicles and the improvement of traffic facilities have brought great convenience to people's travel, but they have also caused a series of traffic accidents. According to relevant statistics, among all kinds of accident deaths, the proportion of traffic accident deaths is 78.5%. Among them, traffic accidents caused by driver fatigue account for about 20% of the total, accounting for more than 40% of serious traffic accidents [1]. Therefore, it is of great practical significance to prevent traffic accidents in advance and reduce casualties and economic losses.

At present, there are several methods to detect driver fatigue: detection based on physiological signals, detection based on vehicle state, and detection based on computer vision [2]. Among them, the method based on computer vision, which analyzes driver information through image processing and pattern recognition technology, has been widely used in the field of fatigue driving monitoring, and has become the most popular at present.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

In visual feature fatigue driving detection, literature [3] uses the face detection algorithm based on Dlib library to locate the eye position, and then uses the eye fatigue parameters to judge whether the driver has fatigue driving. Literature [4] use the traditional face detection algorithm based on VJ(Viola-Jones) is used, and then the eye segmentation is realized according to the color space conversion, so as to further realize the fatigue driving detection. Literature [5] puts forward a face detection method using Haar-like features and Adaboost algorithm to train a strong classification set composed of several weak classifiers. Based on convolutional neural network model, the face of agricultural machinery drivers is divided into eyes closed. The face detection method in literature [6] uses a highly accurate target classifier method based on HOG features and SVM training.

However, in today's society, face detection technology faces many challenges, such as low detection efficiency and inability to accurately locate the target face due to complex backgrounds such as illumination, occlusion and posture. Especially under the influence of mask's large area shielding factor. The face detection used in the above methods is based on prior knowledge, such as facial skin color, facial features, edge and texture features, etc. Another one is based on statistical learning methods. Such as principal component analysis (PCA), SVM, AdaBoost and other machine learning methods. The methods mentioned above have achieved certain results in detecting the driver's fatigue state, but most of them still have a certain scope of application, and further research is needed to extract the features of face state in a complex background. Moreover, the face detection effect will directly affect the recognition of the subsequent modules of the fatigue detection algorithm.

In order to solve the above problems, this paper uses the powerful feature extraction function of neural network, and uses residual network to connect SSD network. The front structure of the network uses residual module, and its network performance far exceeds that of the traditional network model, which can achieve a good face detection effect for mask occlusion. On this basis, Hough transform pupil circle detection algorithm is used to detect the eye position and opening and closing state. The threshold values of PERCLOS (Percentage of eyelid closure over the pupil overtime) and blink frequency are studied, and the experiment and analysis of fatigue driving state detection are completed.

The rest of the paper is organized as follows: The second section explores the related work of face detection network. The third section discusses the characteristics of fatigue judgment. The fourth section introduces the experimental results. Section five summarizes the working methods of this paper.

2. FACE DETECTION ALGORITHM

The research of face detection technology began in 1970s, but it was not until 1990s that it was promoted by computer technology it to made phased progress.

Face detection algorithm is a necessary part of fatigue driving detection. In order to improve the efficiency of face detection, it is necessary to embed an obstacle detection network in the face detection algorithm. At present, face detection and tracking technology is one of the important research directions of target detection. It has been widely used in various applications in daily life, such as video surveillance, human-computer interaction, automatic target tracking, safe driving and medical diagnosis [7-9]. The early research mainly focused on how to correctly detect the face, and didn't think too much about the scene in which the face is located. However, with the expansion of the usage scenarios, from public security to civil economy, from video surveillance to security inspection, people gradually realized that the face detection system should be more general and not only targeted at specific scenes.

At present, the most widely used algorithm is the traditional face detection algorithm based on VJ(Viola-Jones). Its algorithm is mainly aimed at the face images with simple background,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

single scene, similar face size, front-face-oriented, and strong constraints. But in the actual application scenario, the constraint conditions are weak, and the detection accuracy will be affected by many factors, such as scale change, illumination conditions, face angle change, occlusion, blurring, etc., which makes the accuracy of the traditional face detection algorithm drop sharply and can't meet the needs of the actual situation.

In recent years, with the deepening of research, face detection based on Convolutional Neural Network (CNN) has become more and more excellent. With the great success of deep learning in the field of target detection and tracking, its convolutional neural network (CNN) combined with video images also promotes the research of fatigue driving detection. There are also great breakthroughs in CNN's face detection technology, such as MT CNN (Multi-Taskconvolitional Neural Network) [10], FastR-CNN[11], FasterR-CNN, YOLOv3 and other face detection networks. Some of the above target detection frameworks can be used for human face detection only with a little improvement. It can fully extract face features from images in complex scenes and improve the detection accuracy.

Literature [12] compares the detection results of existing MTCNN, YOLOV3 and SSD300 on FDDB data sets through experiments, and good results are obtained. Up to now, the traditional machine learning algorithm of target detection has been gradually replaced by neural network, which has a good performance in both speed and accuracy.

2.1. Residual Network and SSD

SSD network [13] is a one-stage target detection network, which is put forward after YOLO to improve the rough anchor design of YOLO network. Its design idea is mainly multi-scale and multi-aspect ratio dense anchor design and feature pyramid, which directly returns the boundary box and classification probability of the target from the feature graph. SSD detection algorithm is to improve the network VGG16, changing the full connection layer FC7 into convolution Conv7 layer, and adding Conv8, Conv9, Conv10 and Conv11 to the subsequent feature extraction network. The SSD network is shown in Figure 1.

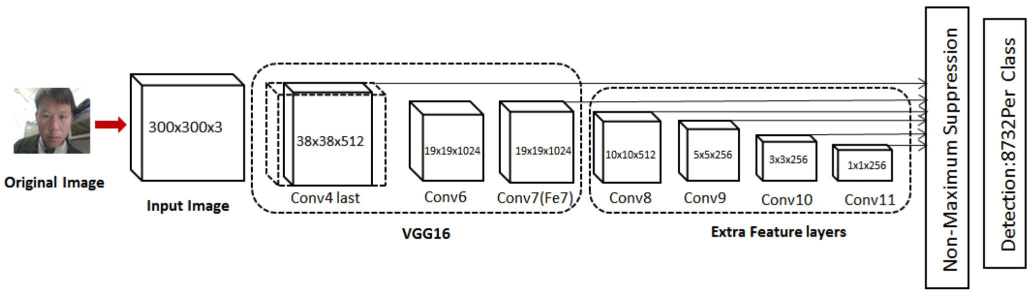

Figure 1. SSD Network Structure Diagram

ResNet(Residual Network) was put forward by Kaiming He in 2015, and achieved very good performance, and won the title of ‘Image Net’ image classification, image object location and image object detection competition in one fell swoop. Res. NET mainly solves the depth problem of convolutional neural network. By adding residual connection in convolutional layer, the performance of convolutional network is improved, which can better deal with image problems. Figure 2 shows the residual connection diagram. Figure3 shows two common models of residual, in which the right side shows the original form of residual module, which is composed of 3x3 convolution with the same output dimension. It is the bottleneck structure of the residual module on the left. The dimension of the input features is reduced by 1x1 convolution layer, the paths are added, and the residual connection is completed.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

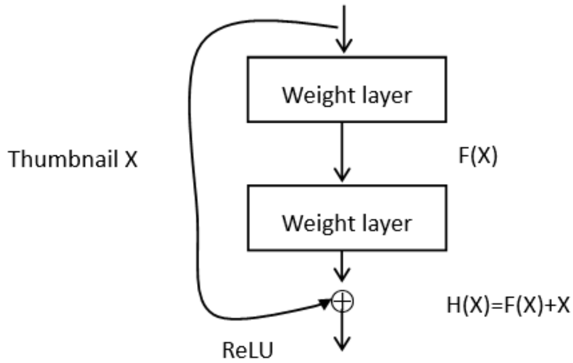

Figure 2. Residual connection diagram

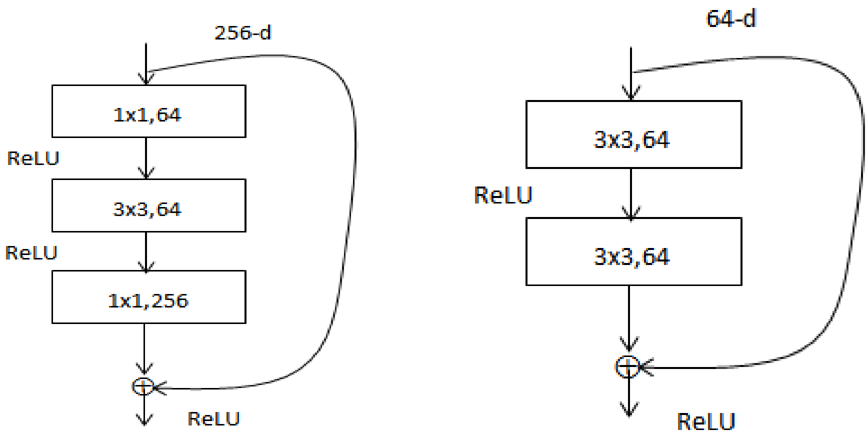

Figure 3. Two forms of residual

Literature [14] achieved good results in face detection by combining ResNet and SSD. Compared with other algorithms, it is shown in Table 1:

Table 1. Comparison of face detection effect of different algorithms in face occlusion [17]

Face detection algorithm	Video frame number	Sample average number of correct frames	Sample average accuracy /%	Frame average detection speed /ms
Haar Adaboost	900	436.4	48.49	67.71
ResNet10 SSD	900	882.0	98.00	19.20
LBP Adaboost	900	125.6	13.96	38.65
HOG SVM	900	202.8	22.53	97.59

The size of the input image is 300x300 pixels, and the number of Residual block is 4. The detection result is shown in Figure 4.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

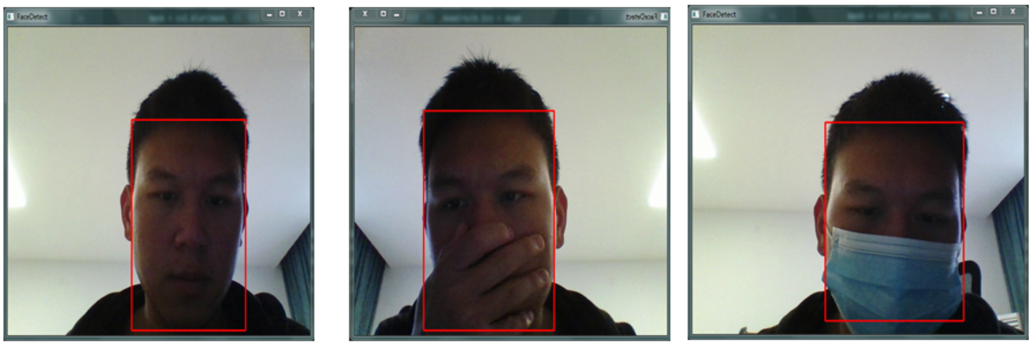

Figure 4. Detection result of face detector

2.2. Pupil detection

Pupil detection, also named eye detection, is the basis of iris recognition and pupil tracking. As a part of face recognition, it has been widely used in many fields. Researchers at home and abroad have done a lot of research on pupil detection. At present, the pupil detection technology is mainly based on image processing and pattern recognition. Table 2 compares the performance and characteristics of different algorithms.

Table 2. Performance and characteristics of pupil detection algorithm

Pupil detection method	Algorithm performance	Algorithm characteristics
Method of measuring characteristics of human body	Low precision and poor robustness	Rough detection in human eye
Skin color model	Low precision and poor robustness	Easy to be affected by light and skin color
Template matching method	High accuracy and general robustness	Not easy to determine standard modules
Method based on time domain information	general precision，general robustness	Processing only for video
Method of machine learning	High precision and accuracy, strong robustness	Need a lot of samples

The eye detection technology based on machine learning has the characteristics of high recognition accuracy and strong robustness, and has become a common method in the field of eye recognition. In this paper, the detection algorithm based on Hough circle transformation is adopted, and the detection and recognition of human eyes can be regarded as the circle existing in the detection image, and it is used as the pupil of human eyes. Hough transform can be effectively used in human eye detection because of its many characteristics. Hough transform has the advantage that it can accurately detect circles in high-quality images, and circles are robust to rotation, generally not affected by noise. Therefore, the research on human eye detection has gradually increased. Figure5 is Hough circle transformation diagram.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

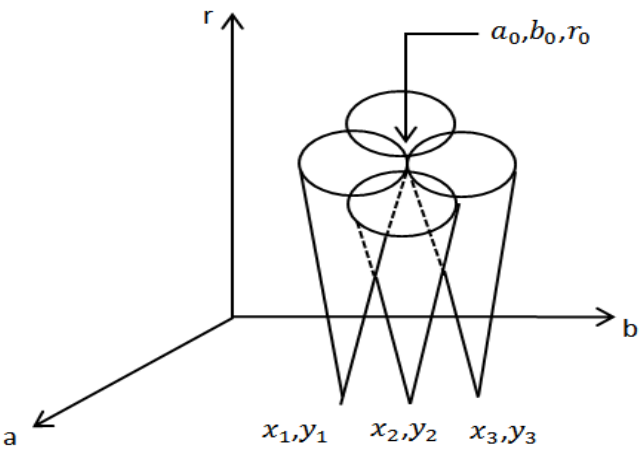

Figure 5. Hough transform schematic diagram

3. FATIGUE FEATURE EXTRACTION AND FATIGUE STATE JUDGMENT

When the driver is fatigued, it is first reflected in his eyes, which can directly show the state of fatigue. When he is awake, the pupil diameter is the largest, and the blinking speed is fast. When he is in severe fatigue, the frequency of blinking is greatly reduced, the aspect ratio of his eyes becomes smaller, the duration of closing his eyes is longer, and the pupil diameter is the smallest. And the eyebrows above the eyes are closely related to the human eyes. When you blink, it will affect the movement of your eyebrows, and when you are tired, it will cause your eyebrows to "rest".

3.1. Blink frequency

Blinking is a physiological action that quickly closes and opens. Everyone blinks a little differently. This mode is different in the speed of closing and opening, the degree of squeezing the eyes and the duration of blinking. Generally speaking, each blinking lasts about 2~4 seconds. In general, our eyes blink about 10-15 times per minute. When it is lower than this number, it can be judged as fatigue state. When the frequency is lower than about 10, it can be judged that the driver is tired. In this paper, we regard the process of eye state from opening to closing and then opening, and the number of closed frames is more than two frames as a blink.

3.2. Eye status judgment

Studies have shown that when people are in a state of fatigue, they will become sensitive to light and narrow their pupils. As in normal situations, the pupils are round, with the same size on both sides. The dynamic characteristics of pupil changes not only indicate the functional status of pupil, but also reflect the status of some human systems and organs. It has broad application prospects in neuroscience, psychology and clinical diagnosis. In this paper, the blinking state can be obtained by detecting the pupil, and the eye-opening state is obtained by detecting the round pupil, otherwise it is the eye-closing state.

3.3. The value of PERCLOS

The test results of driving simulator show that the ratio of eye-closing time can reflect fatigue, as shown in Figure6:

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

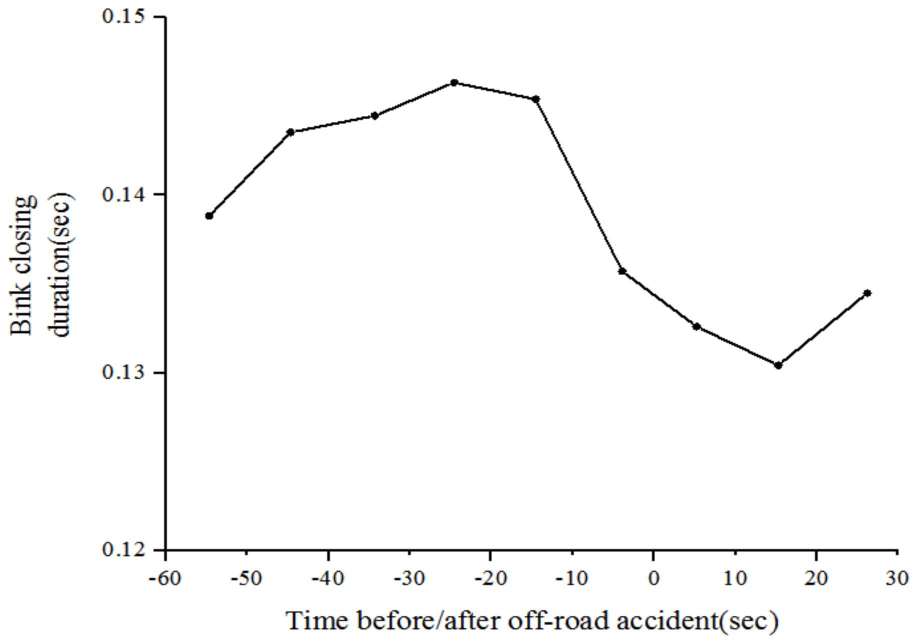

Figure 6. PERCLOS value change diagram

PERCLOS value determines the degree of fatigue by calculating the ratio of eye closing time to a specific time, which is considered to be the best method to measure fatigue at present. Research shows that under normal circumstances, the first eye closing time is between 0.2s and 0.3s. If the eye closing time is longer than the normal eye closing time for several consecutive times, it can be considered as a state of fatigue. There are three methods to determine PERCLOS, namely P70, P80, and EM. The pupil is covered by eyelid by 70%, 80%, and half, respectively. The research shows that P80 has the best correlation with fatigue degree [15], so this paper uses P80 as an index to judge. The formula is PERCLOS=t/T*100%. Where t represents the number of closed frames in time, and t represents the total number of frames per unit time. According to the research [16], when the value of PERCLOS is greater than 0.3, it is judged to be in a state of fatigue.

4. EXPERIMENTAL RESULT

The algorithms used in this paper are all completed in Linux environment, using Python language. In order to determine the blink frequency and PERCLOS threshold suitable for this paper, five segments of 1-minute video of simulated driving in awake state and five segments of fatigue state were taken for testing, and the results were recorded. Table 3 shows the blink frequency and PERCLOS values in awake and tired states.

Table 3. Blinking frequency and PERCLOS value in two states

states	PERCLOS	Blinking frequency
Normal	0.050~0.262	11~30
Fatigue	0.020~0.830	5~14

Randomly select a period of waking state and two periods of fatigue state for calculation. Figure7 shows the test results.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

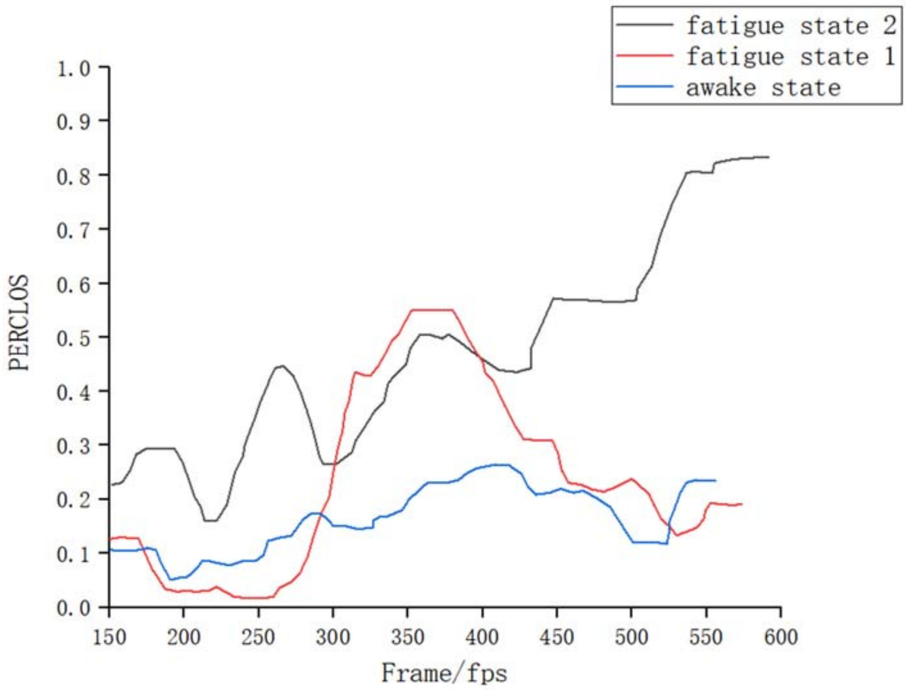

Figure 7. PERCLOS value in three different states

From the above results, it can be seen that the highest value of PERCLOS is about 0.25 in the awake state, and it is obviously larger than that in the normal state in the tired state. Therefore, the threshold of PERCLOS is set to 0.25 in this paper, and when the value of PERCLOS is greater than 0.25, it is judged as fatigue state. By randomly selecting one waking stage and two fatigue stages, blink frequencies in three states are obtained, as shown in Figure8. In fatigue state, blinking frequency is obviously slower than that in awake state. When in a state of fatigue, blink fatigue will become low. In this paper, blink frequency of 10 is selected as the threshold.

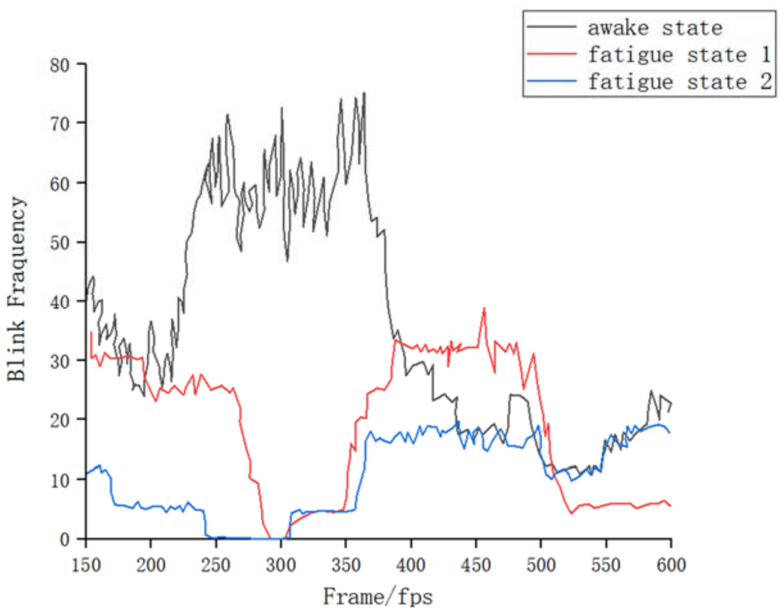

Figure 8. Blinking frequencies in three different states

ISSN: 2472-3703
DOI: 10.6911/WSRJ.202212_8(12).0002

On the basis of face detection, this paper makes an experiment and analysis on fatigue driving detection based on blink detection and pupil detection algorithm. The specific test is shown in the figure, and it can be seen that it has strong robustness.

Figure 9. Testing in natural situations

Figure 9. Testing under mask occlusion situation

Compared with the reference [17], it uses the method of 68 feature points detection on the face of HOG+SVM through Dlib toolbox, and it can be seen that the effect is not good in the case of face occlusion. Can't play a very good detection role. As shown in the figure below.

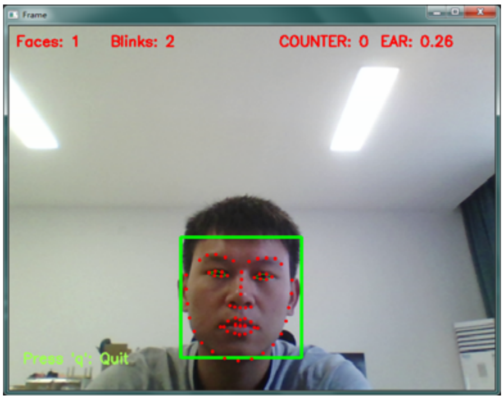

Figure 10. Testing in natural situations

Figure 11. Testing under mask occlusion situation

Five 1-minute videos are collected to test this paper. The video images contain behaviors such as blinking and eye closing in different background environments.

Table 4. Blinking times judgment

Video sequence	Actual Blinks	Algorithm detects blinks
Video 1	6	4
Video 2	10	9
Video 3	15	15
Video 4	20	17
Video 5	9	9

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

Table 5. PERCLOS value judgment

Video sequence	Actual PERCLOS value	Algorithm detects PERCLOS value
Video 1	0.25	0.24
Video 2	0.33	0.32
Video 3	0.15	0.10
Video 4	0.82	0.80
Video 5	0.67	0.68

The above table shows that the blink times and PERCLOS values detected by this algorithm are close to the real values, which further lays a more effective foundation for judging fatigue.

5. CONCLUSIONS

In this paper, the face is roughly located, and then it is finely located to the human eye. Then, whether the driver is in a state of fatigue driving is further determined by the state of human eyes, and the fatigue detection of the driver is carried out by combining PERCLOS and blink frequency. The threshold values of PERCLOS and blink frequency are determined and verified. Compared with other traditional face detectors, this method has good face detection effect, good accuracy of eye indicators and strong anti-interference. The experimental results show that the algorithm can effectively detect the occlusion state, and the detection accuracy and real-time performance are also improved. It lays a better foundation for further fatigue driving detection, hoping that the following work can achieve more feature fusion fatigue detection, and also hope to achieve multiple fatigue grades division; Such as: mild fatigue, moderate fatigue, severe fatigue, wakefulness, etc. It can not only be simply divided into two states, fatigue and wakefulness, fatigue is a gradual process.

REFERENCES

[1] Wang Qiong, Huan Wang, Zhao Chunxia, Yang Jingyu. Driver Fatigue Monitoring based on Eye State Recognition [J]. Journal of Nanjing University of Science and Technology (Natural Science Edition), 2010,34(04):448-453.

[2] L. Song, L.Qiao, "Simulation Research on detection of driver's fatigue," vol. 29(2012), no.11, pp.378-381

[3] Huang Jiacai, Kuang Wenteng, Mao Kuancheng. Research on fatigue driving detection based on face key points [J]. Journal of Nanjing Institute of Technology (Natural Science Edition), 2017,(04):8-13.

[4] Chen Qiang, Dai Jun. Research on fatigue driving behavior of agricultural machinery drivers [J]. China Journal of Agricultural Machinery Chemistry, 2021,42(05):148-152.

[5] Wu Xueying, Wu Caishuo, Huang Wencong, Qin Shulin. Vehicle fatigue driving detection based on Viola-Jones framework face detection algorithm [J]. Journal of Guangxi University of Science and Technology, 2021,32(01):49-54

[6] Zhu Mingliu, Li Dinggen. Research on fatigue detection method based on facial feature points [J]. Computer Application Research, 2020,37(S2):305-307.

[7] Subburaman V B , Marcel S . Alternative search techniques for face detection using location estimation and binary features[J]. Computer Vision & Image Understanding, 2012, 117(5):551-570.

[8] Jie H , Li S , Gang S , et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, PP(99).

[9] Chen J , Li K , Deng Q , et al. Distributed Deep Learning Model for Intelligent Video Surveillance Systems with Edge Computing[J]. IEEE Transactions on Industrial Informatics, 2019:1-1.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202212_8(12).0002

[10] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks.[J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.

[11] Girshick, R.(2015), FAST R-CNN, Proceedings of International Conference on Computer Vision, 1440-1448.

[12] Chen Lili, Zhang Zhengdao, Peng Li. Real-time detection method based on improved SSD [J]. Progress in laser and optoelectronics,2019,56(01):109-115.

[13] Liu W, Anguelov D, ErhanD, et al. SSD:Single Shot MultiBox Detector [C]. European Conference on Computer Vision. Springer, Cham, 2016.

[14] Pan Jiankai, Liu Zhengqing, Wang Qiucheng. Fatigue driving detection based on co-occurrence matrix of eye self-quotient graph-gradient graph[J]. journal of image and graphics, 2021, 26(01): 154-164.

[15] Dinges D F, Grace R . PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance[J]. 1998,Tech Brief.

[16] Wang Lei, Wu Xiaojuan, research progress of detection methods of driving fatigue and drowsiness [J]. Journal of Biomedical Engineering, 2007,24(1):245-248.

[17] Feng Xiangrong, Ma Yukun, Li Linfang, An Jinliang, Zhang Zhixia. A new fatigue driving detection algorithm-NMEAR [J]. Journal of Henan University of Science and Technology (Natural Science Edition), 2021,49(01):65-70+77.