Research on fatigue detection based on visual features

Guangzhe Zhao1 | Yanqing He1 | Hanting Yang1 | Yong Tao2

1 College of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China
2 College of Beihang University, Beijing, China

Abstract
The high incidence of traffic accidents brings immeasurable losses to life. In order to avoid such crises, researchers and automakers have used many methods to solve this problem. Among them, technology based on visual features is widely used in driver fatigue detection. As fatigue detection plays a vital role in the driving process, the high accuracy of fatigue monitoring is very important. This paper focuses on the method based on convolutional neural network to detect driver fatigue. First, in the face detection part, the Single-Shot Multi-Box Detector algorithm is used to improve the speed and accuracy of face detection to extract the eye and mouth regions; second, the VGG16 network is used to learn fatigue features, which is performed on the NTHU-Drowsy Driver Detection (NTHU-DDD) data set and the other two modified data sets Training test. The main result of this work is that the accuracy of fatigue monitoring is higher than other methods including the original method, with an accuracy rate of over 90%. And it has better generalization ability than the multi-physical feature fusion detection method. At the same time, we propose the fatigue detection method based on convolutional neural network to improve the advanced driver assistance system (ADAS) to make it more robust and reliable decision making.

1 INTRODUCTION

The driver is very important in the driving process, and related investigations show that the driver’s lack of concentration is the main cause of traffic accidents. In 2011, one-fifth of the fatal collision research surveys in Canada involved driver fatigue. In 2014, the NHTSA (National Highway Traffic Safety Administration) report showed that there were 846 traffic accidents related to drowsy drivers. In a survey in Pakistan, it was recorded that 34% of traffic safety accidents were caused by driver fatigue [1]. In 2017, a European Union statistics showed that 20% of transportation accidents are related to fatigue driving. Based on these shocking statistics and results, the research community attaches great importance to the driving status of the driver and has conducted a series of studies. The driver’s fatigue state mainly shows behaviours such as closed eyelids, distraction, repeated yawning, regular head movement etc. During the driving behaviour of the driver, there may be more than one symptom or different degrees of fatigue. It is necessary to combine multiple driving behaviours to judge the driving status of the driver.

Fatigue detection technology is mainly divided into subjective testing methods and objective testing methods. The subjective detection method mainly relies on the driver’s subjective judgement of fatigue, and the reliability is low. In contrast, the objective detection method has better accuracy. One of the earliest methods for subjective assessment of fatigue is the Karolinska Sleepiness Scale, KSS [2]. KSS is a subjective assessment questionnaire to obtain the fatigue levels of factory workers and long-distance drivers. KSS describes nine complete fatigue levels, each with an annotation, as shown in Table 1. This is a subjective measurement table that can be used to self-assess fatigue. And the questionnaire data are recorded after a long time, so the driver’s response cannot always be obtained in real time. Therefore, it cannot be used to continuously monitor individual fatigue levels, but KSS provides a benchmark for objective detection methods.

Objective detection methods are the main direction of current fatigue detection research. Objective detection is mainly divided into two categories: fatigue detection based on visual features and fatigue detection based on non-visual features. The detection based on visual features is mainly based on computer vision to extract the physical features of the face such as eyes and mouth to detect whether the driver is fatigued. Building some real-time products, therefore, methods based on visual features have broad research prospects. Visual feature-based
detection methods are divided into four categories in detail: analysis based on eye feature status, analysis based on mouth features, analysis based on head features and fusion analysis of multiple physical features.

Fatigue detection based on non-visual features is mainly divided into two categories: fatigue detection based on vehicle features and fatigue detection based on biological signal features. The fatigue detection method based on vehicle characteristics mainly determines the degree of drowsiness of the driver by measuring the steering wheel movement of the driving vehicle or the distance away from the lane or lateral position. The limitation of this type of technology is that the method relies heavily on specific driving conditions, and the method detection is not flexible enough in essence. Fatigue detection methods based on biosignal characteristics focus on electronic biosignals. They are based on the fact that biosignals start to change in the early stages of driver drowsiness. This method provides an objective and accurate method to detect drowsiness. This method also has limitations. The driver needs to wear various sensors during the process of measuring biological signals, which is invasive to the driver. This makes the application in real life unacceptable.

Based on the important role of visual features in driver fatigue detection, this paper mainly analyses the driver's facial features. Fatigue detection based on driver's facial behaviour is mainly divided into three steps: face detection, feature extraction and fatigue judgement. There are many fatigue detection algorithms based on facial behaviour. In 1994, Wierwille et al. established the percentage of eyelid closure rate over time (PERCLOS) model to assess the degree of fatigue. Through experiments, it is concluded that the role of the PERCLOS model in fatigue detection cannot be ignored. After that, the PERCLOS model has also been widely used [3]. In 2014, Taigman et al. proposed the DeepFace method for face detection, which greatly improves the detection rate compared with traditional face detection methods [4]. In 2016, Mandal et al. proposed a vision-based fatigue detection system, which includes face detection, eye detection and head and shoulder detection [5]. In 2017, the authors performed fatigue detection by calculating the PERCLOS and blinking frequency [6].

This study proposes the current fatigue detection methods, and conducts experimental research on the driver fatigue detection methods based on deep learning. This method has broad application prospects in advanced driver assistance systems (ADAS). Combined with lane detection, obstacle detection and automatic parking, pedestrian detection forms a more complete unmanned driving assistance system, enabling the driver to have a more perfect driving experience.

The main contribution of this paper could be given as follows: (A) We propose a method to detect fatigue based on the fusion of multiple physical features. This method uses Haar-like and 68-Landmark models for face detection, combined with PERCLOS and the aspect ratio of the mouth (MAR) for fatigue detection. (B) We propose a driver fatigue detection method based on convolutional neural network. First, the Single-Shot Multi-Box Detector (SSD) network is used to perform face detection tasks. Then use the VGG16 network to learn the features of the eyes and mouth, and finally perform experiments on the public data set, and obtain good results. (C) We propose a driver fatigue detection system based on convolutional neural networks to complete the advanced driving assistance system, which together with pedestrian detection system, lane detection system, automatic parking system and obstacle detection system form a more comprehensive ADAS.

This paper is divided into six parts. The first part introduces the research background of driver fatigue driving. The second part introduces the definition of driver fatigue and related research progress. The third part introduces the framework of the proposed fatigue detection algorithm. The fourth part introduces the experiment on the expected method and analyses the experimental results. The fifth part proposes the application of fatigue detection. The sixth part summarizes the full text and future prospects.

2 RELATED WORK

In this section, we summarize and introduce the fatigue detection methods. Based on the high accident rate caused by driver fatigue, researchers have proposed various methods to detect fatigue. We list the conventional methods of fatigue detection and their latest research progress, and then the latest methods using deep learning. Before understanding the research progress of fatigue detection, we first briefly introduce the definition of fatigue.

2.1 Definition of fatigue

The American Coordinating Committee for Factors (DOT) provides a detailed description of fatigue. ‘Fatigue is a complex state, lack of alertness, decreased mental and physical function and often accompanied by drowsiness’. Fatigue can be divided into active fatigue, passive fatigue and sleep-related fatigue according to the causes of its operation [7]. Active fatigue is mainly due to a decline in mental and physical functions caused by active participation in tasks. Passive fatigue is caused
Detection method based on vehicle characteristics

This method can detect fatigue through indicators such as lane crossing and steering wheel angle deviation. In addition, the pressure changes on the brakes and accelerators while driving are also powerful indicators of driver fatigue. Research in [13–12] shows that fatigue can be effectively distinguished through vehicle characteristics, and corresponding commercial products have been developed. The detection methods based on vehicle features are mainly divided into steering-based implementation, lane deviation-based implementation, posture-based implementation and multi-vehicle feature fusion methods. This detection technique using lane deviation is highly dependent on driving skills and road conditions. In 2009, Mercedes-Benz designed a system. It first evaluates the driver’s behaviour by monitoring the steering wheel and steering speed. In addition, lane deviation is also widely used in driver fatigue detection systems. Yang et al. [8] designed a driving simulator test bench with a total of 12 volunteers participating. The simulation proved that there is a link between lane deviation and driver fatigue, but this method is only tested on the simulator.

Detection method based on biological signal

This method is based on the fact that at different levels of alertness, physiological signals such as electroencephalogram, electrocardiogram and skin activity will change slightly. Brain signals contain information about brain activity. The signals used to measure fatigue are mainly alpha, delta and theta. When the driver’s level of alertness decreases, the three signals will change to varying degrees, mainly reflected in: the delta and theta signals will suddenly increase, the alpha signal will slightly increase, but the increase is not as high as the first two signals. Research in [13–18] proved that the reliability of detecting fatigue based on biological signals is very high. A study by Chai et al. [13] used automatic regression (AR) to segment the data and extract features from it. A three-layer feedforward Bayesian neural network structure was used to classify fatigue. The test accuracy of this method was 89.7%. It is the best accuracy among the three fatigue detection methods. Although the brain signal is the gold standard for measuring fatigue, the driver needs to wear many sensors when using the brain signal to detect fatigue, which is very disturbing to the driver. At present, there is no method for contactless extraction of brain signals. Another common method based on biological signals is based on ECG signals. Two important parameters for heart signal to detect fatigue are: heart rate (HR) and heart rate variability (HRV). Studies have shown that HRV increases with increasing fatigue. With the development of non-invasive sensors, Jung et al. [15] proposed a method of embedding electrodes into the steering wheel. This method requires high-precision sensors to observe the subtle changes in the driver’s driving state. Secondly, because the sensor is embedded in the steering wheel, it is extremely susceptible to human factors.

Detection method based on facial features

Computer vision-based facial feature extraction method, in which facial features used roughly include eyes, mouth and head. Specific characteristics include the percentage of closed eye time, closed eye time, blink frequency, yawn frequency and nodding frequency. The most suitable and most appropriate indicator for detecting fatigue based on eye condition is the PERCLOS. The studies in [5, 19–22] all show the effectiveness of detecting fatigue based on physical characteristics. There are also studies that use multi-scale parallel convolutional neural network structure or semantic features to recognize facial information and extract features [23, 24]. In recent years, researchers have studied 3D technology and used it to extract more facial parameters [25]. Mandal et al. [5] used PERCLOS to perform fatigue tests on bus drivers. But this method is for the spherical camera on the bus, and is suitable for low-resolution images, and does not have wide applicability. Another widely used method based on facial feature extraction is recognition based on mouth state, and yawning has also proved to be a good indicator of fatigue.

Weng et al. used a new hierarchical time deep belief network (HTDBN) method to identify fatigue by extracting facial features [26]. Jabbar et al. proposed a fatigue detection system based on the multi-layer perceptron classifier based on the key points of facial landmarks, which is mainly designed for Android mobile devices, and the accuracy rate needs to be improved [27]. Jabbar and others pass the captured images to a convolutional neural networks (CNN)-based deep learning model to detect driver fatigue driving behaviour [28].

3 | METHOD

This section introduces two methods of fatigue detection, namely, a detection method based on multi-physical feature fusion and a detection method based on deep learning, and
conducted experiments to prove the feasibility. The operation process diagram of fatigue detection is shown in Figure 1.

3.1 Detection method based on multi-physical feature fusion

Due to the diversity of fatigue characteristics, a single fatigue characteristic is difficult to be used as a standard for fatigue detection. Therefore, the model uses a variety of facial physical feature fusion strategies.

3.1.1 Extract eye features

The face detection algorithm we choose Haar-like features. This method was first proposed by Papageorgiou et al. for face representation[36]. Haar-like features are mainly divided into three categories: edge features, linear features, centre features and diagonal features. There are two kinds of white and black rectangles in the feature template. The feature value in this template is equal to the sum of white rectangular pixels minus black rectangular pixels. In essence, the Haar-like feature value reflects the greyscale change of the image, for example: the eyes are darker than the cheeks, and the mouth is darker than the surrounding colours.

The aspect ratio of the eyes (EAR) can reliably estimate the degree of driver’s eye opening, and plays a vital role in fatigue detection. In [29], the authors described the importance of EAR in detail. First use Dlib’s pre-trained face detector to locate the eye contour according to the output 68-Landmark model. The 68-Landmark model is shown in Figure 2. Second, calculate the EAR to judge whether the eyes are closed. Equation (1) shows the detailed information for calculating the EAR, where \( P_1, P_2, P_3, P_4, P_5 \) and \( P_6 \) are the corresponding six eye feature points in the facial landmark detector. In other words, if the eyes are open, the EAR is basically unchanged, and when the eyes are closed, the EAR approaches zero.

\[
\text{EAR} = \frac{|P_2 - P_4| + |P_3 - P_5|}{2|P_1 - P_6|}. \tag{1}
\]

A common and effective method for the face is the PERCLOS [30, 31], which refers to the percentage of eye closure time in a certain period of time. The fatigue state of the driver depends on the closed state of the eyelids. Many techniques have proven that this technique is very reliable and can be used in conjunction with other facial features related to drowsiness, such as yawning. The calculation of PERCLOS is as in Equation (2). Finally, according to EAR and PERCLOS, we set that if the eyes are closed continuously for more than nine frames, the driver is judged to be fatigued.

\[
\text{PERCLOS}() = \frac{\text{Frames with closed eyelids}}{\text{Total frames}} \times 100%. \tag{2}
\]

3.1.2 Extract mouth features

First, we locate the mouth contour according to the 68-Landmark model. The 68-Landmark model is shown in Figure 2. Second, the geometric method is used to calculate the MAR to judge the state of the mouth. Equation (3) shows detailed information for calculating the MAR, where \( H \) is the height and \( L \) is the width. At the same time, four points of 50, 52, 56 and 58 were selected to define the height of the mouth, and 48 and 64 to define the width of the mouth. By calculating the MAR value, the opening and closing of the mouth can be measured, mainly to detect whether the driver is yawning, and then to judge the degree of fatigue. The calculation of the
specific value of MAR is described in detail in Section 4.2.1.

\[ MAR = \frac{H}{L}. \]  

### 3.2 Detection method based on deep learning

This section introduces the fatigue detection network architecture. The method used in this paper mainly includes two steps: first is the face detection algorithm, and second is the fatigue detection network model.

#### 3.2.1 Face detection

In the face detection part, we chose the Single Shot Multi-Box Detector (SSD) algorithm. The traditional face detection algorithm is based on machine learning such as support vector machines, which uses cascade, AdaBoost and other strategies. It has a small amount of calculation, good real-time performance and poor accuracy.

\[ f(x) = \text{sign}(w \ast x) + b, \]  

The face detection algorithm based on deep learning, due to the huge computational complexity of the complex structure of the convolutional neural network, the accuracy rate is better, but the real-time performance needs to be improved. SSD is a brand-new convolutional neural network target detection algorithm proposed by Liu et al. Its characteristics are: first, feature maps of different scales are extracted for detection, large-scale feature maps are used to detect small objects, and small feature maps are used to detect large objects; Secondly, SSD uses a priori frames with different scales and aspect ratios. It overcomes the shortcomings that the YOLO algorithm is difficult to detect small objects and the positioning is not accurate. This algorithm uses the VGG16 network as the backbone network, replaces the fully connected layers of VGG16 with 3 \times 3 convolutional layers and 1 \times 1 convolutional layers. And the overfitting layer and fc8 layer have been removed, and a new convolutional layer has been added. The network structure of SSD is shown in Figure 3.

In addition, the described convolutional layer also includes activation functions ReLU and Normalization. ReLU is a typical non-linear activation function, and its formula is \( f(x) = \max(0, x) \), which maps the input signal to the feature space. First, compared with the traditional Sigmoid activation function that needs to calculate the exponent and the reciprocal, ReLU has a lower computational cost and faster speed. Secondly, the ReLU function is sparse. That is, it is not activated when the input is less than 0. Compared with the sigmoid activation function with an activation rate of about 50\%, a lower activation rate can be obtained, which is closer to the biological signal transmission process when the brain is working. This has a significant effect on alleviating the overfitting problem. Since Conv4_3 in VGG16 is used to detect the first feature map, the size of the feature map is 38 \times 38. This layer is relatively forward, so a Normalization is added at the back, and each pixel is normalized in the channel dimension to ensure that the difference from the subsequent convolutional layer is not small.

For the optimization algorithm, we use the Momentum algorithm. Momentum algorithm is an improvement of the local minimum point vibration problem in the stochastic gradient descent optimization space. The iterative update formula of the Momentum algorithm is shown in Equation (4). \( \beta \) is an index of gradient accumulation, and we generally set the value to 0.9. \( \alpha \) is the learning rate of the network. \( dw \) is the original gradient calculated by us, and \( r \) is the gradient calculated by exponentially weighted average. This is equivalent to smoothing the original gradient and then using it for gradient descent. When we use the Momentum optimization algorithm, we can solve the problem of excessive swing in the update range of the stochastic gradient descent optimization algorithm, and at the same time make the network convergence faster.

\[ v = \beta v + (1 - \beta)dw, \]  

\[ r = \beta r + (1 - \beta)dw. \]
3.2.2 VGG16 convolutional neural network

We extract frames from the video, these frame images are fed to the SSD network for face detection, and then, the detected eyes and mouth in the face are used for fatigue feature learning. We extract frames from the video, these frame images are fed to the SSD network for face detection, and then the detected facial eyes and mouth images are used for fatigue feature learning.

VGG16 is a simple convolutional neural network. It consists of several $3 \times 3$ convolutional layers and $2 \times 2$ pooling layers repeatedly stacked. VGG16 contains a total of 13 convolutional layers, 3 fully connected layers and 5 pooling layers.

The convolutional layer and pooling layer of VGG16 can be divided into different blocks, in order from block 1 to block 5. Each block contains several convolutional layers and one pooling layer, and, in the same block, the number of channels of the convolutional layer is the same. For example, the second block contains two convolutional layers, and each convolutional layer is represented by $3 \times 3 \times 128$, that is, the convolution kernel size is $3 \times 3$, and the number of channels is 128. The network convolution structure of VGG16 is shown as in Figure 4.

4 EXPERIMENT

In this section, we propose the experimental data and results. At the same time, the accuracy between different methods is proposed in detail.

4.1 Data set

The NTHU-Drowsy Driver Detection (NTHU-DDD) video data set is a data set proposed by Weng et al. [26] and has been extensively studied. This data set consists of five scenes without glasses, with glasses, with sunglasses, without glasses at night and with glasses at night, and is composed of male and female drivers of different races. The video is $640 \times 480$ pixels, 30 frames per second, and each frame is marked as sleepy or non-sleepy. The sample shown in Figures 5–9 shows the richness of the data set.

4.2 Result

4.2.1 Multi-physical feature fusion detection method

First of all, in the process of extracting the features of the mouth, how to make the distinction between yawning and speaking becomes very important. Through experimental research, the four states of mouth closing, speaking, smiling and yawning were tested. The experimental results are shown in Table 2.

| Mouth state | Height $h$/mm | Width $l$/mm | Aspect ratio MAR |
|-------------|--------------|--------------|-----------------|
| To shut up  | 32           | 60           | 0.5333          |
| Speak       | 25           | 57           | 0.4386          |
| Smile       | 23           | 65           | 0.3538          |
| Yawn        | 61           | 54           | 1.1296          |
Analysis of Table 2 shows that in the yawn state, the MAR is 1.1296, and the other three cases are less than 1. Ten more videos of yawning in fatigue driving state were extracted, compared with MAR in speaking state and finally the threshold value in yawning state was determined to be 0.85. In addition, a yawn judges the driver’s fatigue is inaccurate. Therefore, we have designed a yawn counter. For example, if there are more than three yawns in a certain period of time, it will be judged as fatigue.

Second, the setting of PERCLOS threshold is also very important, which will affect the accuracy of fatigue detection. Studies have shown that the normal blink time of a person is 0.2–0.3 s, and the value of PERCLOS is between 16.7% and 100%. We set different PERCLOS values to calculate on the data set, and finally select the value with the highest fault tolerance rate and set the PERCLOS threshold to 20%.

As described in Section 3.1, the model of multi-physical feature fusion is described. In order to verify the performance of the model, the model was tested under various conditions. 300 photos were input to the system for the test set, with an accuracy rate of 95.7%. At the same time, we use Opencv to call the camera to simulate the driving environment for verification. Real-time display of eye aspect ratio (EAR) and MAR as well as
4.2.2 Detection method based on deep learning

In Section 3.2, the fatigue detection method based on VGG16 network is described. We have verified on two data sets separately. The accuracy rate of VGG16 verification in the homemade data set is 91.4%. In the NTHU-DDD data set, the verification accuracy of VGG16 is 91.88%. Table 3 shows the results of our overall experiment.

We compared the VGG16 network used with other methods using the NTHU-DDD data set for experiments, especially the original method for publishing this data set [26]. Mainly in three aspects of training methods, models, verification accuracy and so on. The original work used a new layered time deep trust network (HTDBN) as the training model. According to the description, the verification accuracy of this method is about 85%. In [27], the method of feedforward neural network is used to verify the NTHU-DDD data set, and the accuracy rate is 80%. In [28], the accuracy of drowsiness detection based on CNN and facial landmark detection (D2CNN-FLD) is 83.3%. At the same time, experiments are carried out based on the AlexNet convolutional neural network and RCNN convolutional neural network. The accuracy of the AlexNet convolutional neural network is 82.8%, and the accuracy of the RCNN convolutional neural network is 90.5%. In contrast, the verification accuracy of the VGG16 network is 91.88%, both exceeding other methods. The accuracy comparison between our method and the latest method is shown in Table 4.

5 APPLICATION

ADAS is one of the current research hotspots in automotive intelligence. Its perception system is its core component. The perception system detects and recognizes people, vehicles and objects, and provides support for vehicle decision making. In recent years, although the performance of automobiles has been developing rapidly, there are still many major traffic accidents caused by driver fatigue. Traffic accidents cause intangible losses to people, and avoiding driving when the driver is fatigued is an important way to prevent such accidents. The government has legalized some policies to reduce traffic accidents. For example, when a driver drives for more than 4 h in a row, a mandatory break is required. But these preventive measures are not enough to reduce the incidence of traffic accidents. Although the development of ADAS is becoming more and more mature, and corresponding commercial products have been developed. However, there is still a long distance from unmanned driving. In the future, automatic driving is still in the assisted driving stage, that is, the driver's participation is required.

Over the past decade, the automotive industry has developed systems for driver fatigue detection. Companies including Ford, Volkswagen, Toyota etc., have developed various technologies on their products to detect driver fatigue, to improve ADAS, and assist driver driving. In 2006, Toyota first developed a driver monitoring system and proposed an advanced obstacle detection system. That is, during driving, the driver is detected to turn his head away from the road, and an obstacle is detected on the road, the system will warn the driver. In 2007, based on vehicle characteristics information, Volvo combined the driver warning control system and lane departure system to introduce a driver drowsiness detection system. In 2012, Volkswagen launched a driver fatigue detection system that can automatically analyse driving characteristics.

In addition, business solutions focusing on the use of smartphones have been proposed, which fully demonstrate the applicability of mobile technology to assist in safe driving. Among them, the video frames obtained by the mobile application based on the Android smartphone combined with the computer vision image processing technology detect the driver's drowsiness. After experimental testing, the accuracy rate is 85% [27]. However, this system has limitations for rapidly changing lighting and sunglasses. With the development of technology, more and more portable and fashionable wearable products appear in life. Using their non-intrusiveness to overcome the shortcomings of traditional biometric detection, it is wise to effectively use them in the driver fatigue detection system.

| Table 3 | Summary of overall experiment results |
|---------|--------------------------------------|
| Multi-physical feature fusion method | VGG16 DATA SET1 | VGG16 NHTU-DDD |
| Accuracy (%) | 95.7 | 91.4 | 91.88 |

| Table 4 | Comparison of results between our method and the state-of-the-art methods |
|---------|--------------------------------------|
| Research | Year | Method | Accuracy (%) |
| Weng [26] | 2016 | HTDBN | 84.82 |
| Shih [33] | 2016 | MSTN | 82.61 |
| Jabbar [27] | 2018 | MLP | 80.92 |
| Xing [34] | 2019 | CNN+PERCLOS | 87.5 |
| Li [35] | 2019 | Multi-feature fusion | 86.25 |
| Jabbar [28] | 2020 | D2CNN-FLD | 83.3 |
| Jabbar [28] | 2020 | AlexNet | 82.8 |
| Jabbar [28] | 2020 | Faster RCNN | 90.5 |
| Ours | - | SSD+VGG16 | 91.88 |
ADAS mainly includes lane detection system, obstacle detection system, automatic parking system and pedestrian detection system. We believe that a more complete ADAS should include both inside and outside warning mechanisms. All single auxiliary methods have certain limitations in actual use scenarios. All available commercial systems and non-commercial solutions should be regarded as auxiliary systems, the purpose of which is to assist the driver in driving and avoid traffic accidents. We believe that the summarization of the outputs of different sensors (such as cameras, vehicle sensors and human sensors) will ultimately lead to more robust and reliable decisions. One possible solution for aggregation is to assign weights to each solution based on its flexibility. All in all, these systems are in progress.

6 CONCLUSION AND FUTURE WORK

This paper uses the method based on physical feature fusion and the method based on deep learning, respectively, to detect driver fatigue. This paper focuses on the method of fatigue detection based on convolutional neural network. The experimental results in various situations provide the possibility for the realization of the driver's fatigue detection system. We perform fatigue classification based on the faces detected by the SSD network. The results show that eyes and mouth are important features that play an important role in fatigue detection. The method of combining eyes and mouth achieves 95% accuracy. At the same time, we made the NTHU-DDD data set. The detection method based on the VGG16 network has 91.88% accuracy on this data set, which is about 5% higher than the original method. Experiments prove that our method has better accuracy. The result of this work is that it can be used in ADAS for high-precision, high-safety driver fatigue detection, and has a wide range of applications in ADAS.

In future work, 3D face information has become a new and inevitable subject. More useful features can be extracted from 3D faces to identify fatigue. At the same time, real-time performance also plays a vital role in fatigue detection. Combine the dense network with knowledge extraction technology, and compare the detection efficiency with the latest methods for real-time application. In addition, consider applying time-series information to fatigue detection.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant 61871021 and the Pyramid Talent Training Project of BUCEA.

REFERENCES

1. Azam, K., et al.: Comparison of fatigue related road traffic crashes on the national highways and motorways in Pakistan. J. Eng. Appl. Sci. 33(2), 47–54 (2014)
2. Kaida, K., et al.: Validation of the Karolinska Sleepiness Scale against performance and EEG variables. Clin. Neurophysiol. 117(7), 1574–1581 (2006)
3. Wieverweil, W.C., Ellsworth, L.A.: Evaluation of driver drowsiness by trained raters. Accid. Anal. Prev. 26(5), 571–581 (1994)
4. Taigman, Y., et al.: DeepFace: Closing the gap to human-level performance in face verification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1701–1708 (2014)
5. Mandal, B., et al.: Towards detection of bus driver fatigue based on robust visual analysis of eye state. IEEE Trans. Intell. Transp. Syst. 18(3), 545–557 (2016)
6. Zhang, F., et al.: Driver fatigue detection based on eye state recognition. In: International Conference on Machine Vision and Information Technology (CMVIT), pp. 105–110.IEEE (2017)
7. May, J.E., Baldwin, C.L.: Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transp. Res. Part F Psychol. Behav. 12(3), 218–224 (2009)
8. Yang, J.H., et al.: Detection of driver fatigue caused by sleep deprivation. IEEE Trans. Syst. Man Cybern. Part A Syst. Humans 39(4), 694–705 (2009)
9. McDonald, A.D., et al.: Real-time detection of drowsiness related lane departures using steering wheel angle. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 56, pp. 2201–2205.SAGE Publications, Los Angeles, CA (2012)
10. Li, Z., et al.: Online detection of driver fatigue using steering wheel angles for real driving conditions. Sensors 17(3), 495 (2017)
11. Furugori, S., et al.: Estimation of driver fatigue by pressure distribution on seat in long term driving. Rev. Automot. Eng. 26(1), 53–58 (2005)
12. Wakita, T., et al.: Driver identification using driving behavior signals. IEICE Trans. Inf. Syst. 89(3), 1188–1194 (2006)
13. Chai, R.: Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system. IEEE J. Biomed. Health Inf. 21(3), 715–724 (2016)
14. Borghini, G., et al.: Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6442–6445.IEEE (2012)
15. Jung, S.J., Shin, H.S., Chung, W.Y., et al.: Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. IET Intel. Transport Syst. 8(1), 43–50 (2014)
16. Balasubramanian, V., Adalarasu, K.: EMG-based analysis of change in muscle activity during simulated driving. J. Bodyw. Mov. Ther. 11(2), 151–158 (2007)
17. Zhu, X., et al.: EOG-based drowsiness detection using convolutional neural networks. In: International Joint Conference on Neural Networks (IJCNN), pp. 128–134.IEEE (2014)
18. Sun, Y., Yu, X.: An innovative nonintrusive driver assistance system for vital signal monitoring. IEEE J. Biomed. Health Inf. 18(6), 1932–1939 (2014)
19. Sigari, M.H., Fathy, M., Soryani, M.: A driver face monitoring system for fatigue and distraction detection. Int. J. Veh. Technol. 2013, 1–11 (2013)
20. Alioua, N., Amine, A., Rziza, M.: Driver’s fatigue detection based on yawning extraction. Int. J. Veh. Technol. 2014, 1–7 (2014)
21. Wang, L., et al.: Gabor-based facial image sequential pattern mining for human fatigue monitoring. Int. J. Adv. Comput. Technol. 3(6), 26–33 (2011)
22. Bergasa, L.M., et al.: Real-time system for monitoring driver vigilance. IEEE Trans. Intell. Transp. Syst. 7(1), 63–77 (2006)
23. Lee, S., Kim, I.: Video captioning with visual and semantic features. J. Inf. Process. Syst. 14(6), 1318–1330 (2018)
24. Li, C., et al.: A multi-scale parallel convolutional neural network based intelligent human identification using face information. J. Inf. Process. Syst. 14(6), 1494–1507 (2018)
25. Zhou, S., Xiao, S.: 3D face recognition: A survey. Human-Centric Comput. Inf. Sci. 8(1), 351–37 (2018)
26. Weng, C.H., Lai, Y.H., Lai, S.H.: Driver drowsiness detection via a hierarchical temporal deep belief network. In: Asian Conference on Computer Vision, pp. 117–133.Springer (2016)
27. Jabbar, R., et al.: Real-time driver drowsiness detection for android application using deep neural networks techniques. Procedia Comput. Sci. 130, 400–407 (2018)
28. Jabbar, R., et al.: Driver drowsiness detection model using convolutional neural networks techniques for android application. In: IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), pp. 237–242. IEEE (2020)
29. Cech, J., Soukupova, T.: Eye blink detection using facial landmarks. In: 21st Computer Vision Winter Workshop, pp. 1–8 (2016)
30. Flores, M.J., Armingol, J.M., de la Escalera, A.: Real-time warning system for driver drowsiness detection using visual information. J. Intell. Rob. Syst. 59(2), 103–125 (2010)
31. Sun, C., et al.: Real-time driver fatigue detection based on eye state recognition. Appl. Mech. Mater. 457, 944–952 (2014)
32. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2(3), 1–27 (2011)
33. Shih, T.H., Hsu, C.T.: MSTN: Multistage spatial-temporal network for driver drowsiness detection. In: Asian Conference on Computer Vision, pp. 146–153. Springer (2016)
34. Xing, J., et al.: Application of face recognition based on CNN in fatigue driving detection. In: Proceedings of the International Conference on Artificial Intelligence and Advanced Manufacturing, pp. 1–5 (2019)
35. Li, X., et al.: Fatigue driving detection model based on multi-feature fusion and semi-supervised active learning. IET Intel. Transport Syst. 13(9), 1401–1409 (2019)
36. Papageorgiou, C.P., Oren, M., Poggio, T.A. General framework for object detection. Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), (555–562.1998)

How to cite this article: Zhao G, He Y, Yang H, Tao Y. Research on fatigue detection based on visual features. IET Image Process. 2021;1–10. https://doi.org/10.1049/ipr2.12207