Driver fatigue detection method based on facial features using deep learning

Wanning Li

School of Computer Science and Cyber Security (Oxford Brookes College), Chengdu University of Technology College, Chengdu, China

li.wanning@student.zy.cdut.edu.cn

Abstract. Fatigue driving can result in delayed driver reactions and lack of concentration, increasing the risk of traffic accidents. Therefore, finding effective fatigue detection methods is crucial for accident prevention and ensuring public safety. Traditional driver fatigue detection methods based on facial features are limited by their reliance on manually designed features and classifiers, leading to low accuracy, slow detection speed, and inefficient performance. Hence, methods based on deep learning have gradually become a trend in fatigue-driving detection research. This paper focuses on driver fatigue detection methods based on facial features using deep learning. In this study, we analyze numerous research outcomes on how to accurately extract, analyze, and integrate various facial features, such as eye features, mouth features, facial expressions, and head postures, to precisely determine the driver's fatigue state. Then, we compare the latest deep learning-based technologies. This paper provides an information-rich repository to help researchers select more promising designs, that can drive the widespread development of fatigue-driving detection technologies.

Keywords: Driver Behavior Monitoring, Deep Learning, Feature Extraction, Facial State Discrimination.

1. Introduction

Traffic accidents caused by driver error pose a significant threat to public safety, so there is an urgent need to take effective measures to reduce the consequences of such incidents. Implementing a reliable driver behavior monitoring system is considered one of the key strategies to reduce driver error [1]. According to a 2015 report by the US National Highway Traffic Safety Administration (NHTSA), 94% of crashes were caused by driver error. Distracted driving and fatigue were found to be the most significant factors, resulting in 41% of crash-related injuries [2].

Driver behavior monitoring plays an important role in ensuring road safety and improving traffic efficiency. By detecting the driver's potentially dangerous driving behavior, it can take preventive measures in time, which can reduce the accident rate. In addition, the detection system can aggregate and analyse information about road use and accidents, providing important information for road planning. This can improve traffic congestion management and optimize the use of resources. Therefore, effective real-time monitoring of driver behavior can help prevent accidents, and improve drivers' bad driving habits, and road safety management.

Fatigue detection can be performed from four aspects: vehicle behavior, driver behavior, driver physiological characteristics, and facial features. Fatigue detection based on vehicle behavior does not require direct contact with the driver or other equipment. It can indirectly assess the driver's driving state by analyzing the vehicle's motion characteristics, including speed, acceleration, lateral sway angle, and lane departure [3]. Fatigue detection based on driver behavior is characterized by easy installation, convenient result acquisition, and low cost. However, it is still affected by factors such as individual driving habits, different types of vehicles, and tires [4]. In addition, fatigue detection that uses driver physiological characteristics typically requires the installation of hardware devices on the driver. Unlike natural driving conditions, wearing complex equipment during testing may impact driver operation [5].

Fatigue driving detection method based on facial features has garnered significant attention due to its human-centric and high-accuracy advantages, becoming the current research focus. Despite the rich research outcomes in this field, most methods exhibit high similarity, slow progress, and a lack of groundbreaking innovation [6]. Traditional driver fatigue detection using machine vision has limitations: reliance on manual feature design hampers adaptability to diverse facial expressions; sensitivity to environmental conditions affects performance in complex scenarios; difficulty in accommodating individual differences weakens generalization; and limited modeling capabilities struggle with dynamic changes in fatigue states. These constraints hinder broader applications. With the rise of deep learning, methods based on deep learning have gradually become a trend to overcome these shortcomings. The deep learning model has automatic learning features and can model complex relationships, which improves the performance and robustness of the algorithm.

Research on the topic of driver signature detection is extensive and straightforward, and several review papers on fatigue detection have been published in the past few years. However, all of these methods are still mainly summarized in traditional fatigue detection methods, such as the review of fatigue detection based on facial features and the systematic review of fatigue detection [6, 7]. There seems to be a lack of more targeted and comprehensive summaries for the research about driver feature detection based on deep learning. Therefore, this paper aims to deeply analyze and compare the current drowsy driving detection methods based on deep learning drivers' facial features. By dividing facial features into several main categories, this paper focuses on the latest deep learning systems, algorithms, and technologies using facial features for detecting driver drowsy drivers. In addition, a detailed comparative analysis of the latest technologies is also carried out, which is presented in the form of tables and other forms. Therefore, this review is a treasure trove of information, not only to introduce researchers in the field to the latest relevant technologies but also to help them select designs that will drive the broader development of drowsy driving detection technologies.

2. Related Work

2.1. Fatigue driving

Fatigue driving refers to a behavior in which drivers experience symptoms such as mental fatigue, distraction, and slow reaction when driving for a long time, lacking adequate rest, or being in a poor physiological state, which increases the risk of traffic accidents. Fatigue driving may associated with factors such as prolonged driving, insufficient sleep, irregular sleep patterns, and continuous work. In the fatigue state, drivers usually face problems such as inattention, blurred vision, and delayed reaction time, which seriously affects driving safety.

The detection of driver fatigue is primarily classified into active detection and passive detection methods. Active detection involves methods such as drivers filling out questionnaires or interviews, relying on subjective awareness. In contrast, passive detection is not influenced by the driver's subjective awareness, offering higher accuracy and reliability. Consequently, passive detection methods represent the primary focus of current research on fatigue driving. Passive methods are mainly divided into fatigue detection based on vehicle behavior, driver behavior, physiological characteristics, and facial features [8].

2.2. Facial fatigue features

2.2.1. Eye features. Eye characteristics during drowsy driving include blink frequency, duration of eye closure, gaze stability, changes in pupil diameter, and yawning. Blink frequency is a commonly used indicator, where an increased blink frequency in fatigue reflects signs of visual fatigue. Additionally, the duration of eye closure is a key feature, as prolonged eye closure beyond the normal fatigue state is indicative of potential fatigue, Then, yawning is usually accompanied by a blinking motion or a long period of eye closure.

2.2.2. Mouth features. In a state of fatigue, there will also be subtle changes in the driver's mouth. Typically, the mouth tends to open more easily due to muscle relaxation caused by fatigue. In simple terms, overall muscle fatigue may be accompanied by a slight tremor in the lips.

2.2.3. Head postures. Under normal circumstances, drivers often look left and right to observe the road conditions regularly. However, in a state of fatigue, the driver may unconsciously exhibit frequent nodding or head-up movements as they try to stay awake and prevent falling asleep. In addition, fatigue can also affect the entire position of the head. For example, fatigue can cause the overall head position to tilt down excessively.

2.2.4. Facial expressions. Focused on the driver's facial movements, facial expressions involve mouth actions, smile intensity, yawning frequency, and overall changes in facial muscle tension. Such as, during fatigue, drivers may exhibit frequent yawning and mouth-opening movements, while the degree of smiling may decrease. These changes in facial expressions offer intuitive reflections of the driver's drowsy state.

2.3. The fatigue driving detection process based on facial features

The process of detecting fatigue driving through facial features primarily involves the utilization of in-vehicle cameras and computer vision technology to capture the driver's facial characteristics. This process includes assessing the driver's facial feature states to determine if they are experiencing fatigue while driving. Figure 1 illustrates the basic workflow of the facial feature-based approach to fatigue driving detection process.

As depicted in the flowchart of Figure 1, the process is initiated by obtaining a data set of the driver's behavior through in-vehicle cameras. After the camera acquires the image, it is fed into the face detection model. Upon detecting the driver's face, facial features are extracted and standardized for subsequent facial state discrimination. Finally, based on the determined facial state, the system assesses fatigue, and if potential signs of fatigue driving are detected, corresponding alert mechanisms are triggered to prompt the driver to take necessary actions, such as resting or changing drivers.

The main differences between traditional and deep learning-based fatigue detection are feature extraction, facial state recognition, and fatigue state discrimination. Compared with the traditional method involving manual design, the deep learning model improves the performance and robustness of the facial state recognition and fatigue detection system through automatic feature learning, multi-feature fusion, end-to-end learning, and complex relationship modeling.
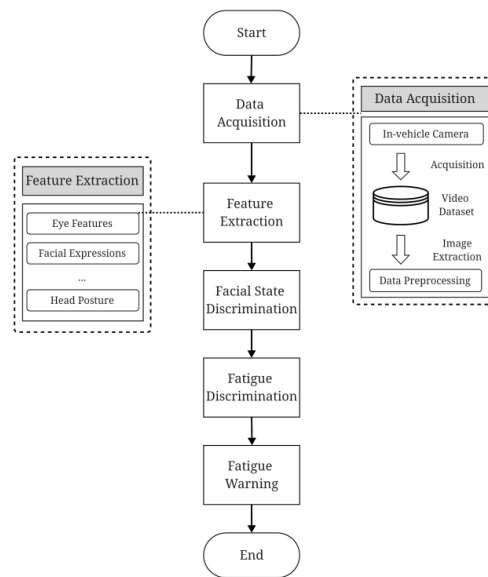
Figure 1. Fatigue driving detection process based on facial features (Photo credit: Original)

3. Fatigue driving detection based on deep learning

3.1. Common deep learning network structures

3.1.1. Convolutional neural network. A convolutional neural network (CNN) is a kind of neural network model specially used to process gridded data in deep learning, it is composed of the input layer, a convolutional layer, a pooling layer, and a fully connected layer, which has become a research hotspot.

The history of CNN can be traced back to 1998 when Professor Lecun introduced LeNet-5, a two-dimensional CNN specifically designed for image applications. It represented a milestone by introducing the concepts of convolution and pooling layers to the field of neural networks [9]. In 2012, Krizhevsky et al. first proposed the Alex-Net network. To solve the problems of hardware limitations, applicability across diverse image categories, overfitting, and vanishing gradients, the network is designed with 5 convolutional layers and 3 fully connected layers [10]. In 2013, Zeiler and Fergus proposed a method to visualize the internal features of CNNs by showing the activation map of the middle layer of the network to help people better understand how CNNs work [11]. In 2014, the Visual Geometry Group (VGG) model was proposed by the Visual Geometry Group. VGGNet has the characteristics of a small convolutional kernel, simplicity, and deep network structure, which improves the ability to locate problems and classify images. However, VGG utilizes approximately 1.4 billion parameters, leading to issues such as excessively high computational costs [12]. Still in 2014, GoogLeNet (also called Inception-V1), proposed by Google, introduced the "Inception" module. GoogLeNet increased the depth and width of the network while maintaining lower computational costs, achieving higher levels of accuracy. However, the heterogeneous topology of GoogLeNet comes with drawbacks such as complexity and difficulty in interpretation [13]. Residual Network (ResNet) was introduced in 2015 by Microsoft researchers. ResNet introduced residual connections to solve the problems of gradient vanishing and exploding in deep networks, which allows the training of even deeper networks. Compared to VGG, ResNet achieves greater depth with lower computational complexity [14].

3.1.2. Recurrent neural network. Recurrent Neural Network (RNN) differs from the previously mentioned feedforward neural networks. It is mainly used to handle sequence data and tasks related to time series, such as speech recognition, time series prediction, and other fields, where contextual information and time series dependencies need to be considered. The core feature of RNNs is the

introduction of a recurrent structure, which allows RNNs to consider contextual information when processing sequential data [15].

Some researchers have found that traditional RNNs have serious gradient explosion and gradient vanishing problems, which resulted in the inability to train deep RNNs. In 1997, Hochreiter and Schmidhuber introduced Long Short-Term Memory (LSTM), a powerful recursive neural system. The memory cell structure in the LSTM allows for better control over the flow of information and can efficiently handle long-term dependencies on sequential data. Compared to conventional RNNs, LSTMs successfully overcome the limitations associated with processing long sequences [16].

3.2. Fatigue detection based on facial features

3.2.1. Face recognition. In the process of traditional fatigue detection, face recognition is an important preparatory step for subsequent face feature extraction and fatigue detection. The common traditional face recognition algorithms include Haar + AdaBoost, based on the OpenCV library, and HOG + SVM, based on the Dlib library [6].

The deep learning model can automatically learn higher-level, abstract features through the training stage. This gives Deep Learning an advantage in complex scenes, such as lighting changes, occlusion, etc. Nowadays, the most widely used face detection algorithms based on deep learning include the Multi-task Cascaded Convolutional Networks (MTCNN) algorithm, Region-based Convolutional Networks (R-CNN) series, Single Shot Multi-box Detector (SSD), RetinaNet, and You Only Look Once (YOLO) series. MTCNN is more flexible than traditional methods because it can automatically learn the features that face recognition need by training on large amounts of labeled data. At the same time, the cascading structure of MTCNN can effectively handle multiple tasks such as candidate image generation, refinement, and key point detection, which improves the overall performance [3]. Furthermore, the R-CNN series, it has improved speed while maintaining accuracy by introducing technologies such as the Regional Suggestion Network (RPN). Besides, the characteristic of SSD is that it uses a single forward propagation method to simultaneously handle targets at the same time. The advantage of RetinaNet is that it has excellent detection speed and accuracy by combining a feature pyramid network and focus loss. Then, the YOLO series is characterized by improving accuracy while maintaining speed by turning the object detection tasks into regression problems. Considering that drivers will have various special situations in real life, such as large posture changes and wearing masks, which affect the accuracy of feature detection, some researchers proposed using the improved YOLOv5 model to improve the detection accuracy of the eyes, mouth, and face [17].

3.2.2. Eye feature. Among the detection methods based on eye features, Wierwille et al. initiated the exploration of the correlation between eye behavior and driver fatigue in 1994 through a series of experiments conducted on a driving simulator. The results of Their research revealed a positive correlation between the frequency and duration of eye closure per unit time and the level of driver fatigue [18]. Building upon this insight, Carnegie Mellon University introduced the Percentage of Eye Closure (PERCLOS) as a key parameter to assess driver fatigue. PERCLOS is computed by analysis of the driver's eyes being open and closed, specifically determining the proportion of time the eyelids remain closed. During extended driving sessions or periods of fatigue, drivers tend to exhibit more frequent and prolonged eye closures. Monitoring the percentage of time these closures occur enables the system to accurately detect signs of driver fatigue [19].

With the continuous development of computer technology, many fatigue detection systems have been proposed based on the changes in the driver's eye condition. The whole process is mainly divided into two parts: Eye detection and fatigue detection. Eye detection can be done using traditional methods based on geometry, feature point calibration, or deep learning networks for eye detection. In the fatigue detection phase, several eye-related facial features are associated with fatigue, including blink frequency, eye closure duration, proportion of prolonged closed blinks (PLCDB), PERCLOS, and yawning [20]. In addition, some researchers have begun to study how to use the eye image directly from the camera,

considering that the eye tracking system is too expensive. Quddus et al. used RNNs to capture eye-related motion in each frame of the video. He used two types of LSTMs, including 1D LSTM (R-LSTM) and convolutional LSTM (C-LSTM). Finally, in a comparison of the results, C-LSTM performed better [21].

3.2.3. Mouth feature. In the current field of driver fatigue detection, mouth characteristics are also one of the important features of driver fatigue state. A large number of studies are beginning to increase the focus on mouth features and supplement them with fatigue detection systems to analyze alongside facial expressions and eye behavior, which make up a complete picture of the driver's overall alertness level.

Many studies have begun to combine eye and mouth features for fatigue detection. Zhao et al. proposed a convolutional neural network EM-CNN to recognise the state of the eyes and mouth. EM-CNN mainly uses PERCLOS and mouth opening (POM) to recognise fatigue based on these two parameters. It is also more suitable for real-world, complex driving conditions with factors such as lighting, seating position, and eye occlusion [22]. Moreover, in the study of Ping et al., a modified MTCNN was used to recognise facial areas, and a lightweight AlexNet classification was used to recognise the states of eyes and mouth. By adjusting the parameters and using the Ghost module, the AlexNet model achieved an accuracy rate of 99.4% while guaranteeing a smaller model size and a short average recognition time. Fatigue was then detected using thresholds for the PERCLOS and PMOT parameters with a fatigue detection accuracy of 93.5% [23]. In conclusion, the integration of mouth feature detection increases the system's ability to recognise potential signs of drowsiness and improves the effectiveness of multimodal approaches to road safety.

3.2.4. Head posture state. The head posture feature provides a nuanced view of the driver's alertness and potential signs of fatigue. Head posture, including movements such as nodding and the overall position of the head, is an important indicator of the driver's cognitive state. Head posture states are as much of a supplement as mouth features. Fatigue detection systems incorporate other facial features, such as head posture, to more accurately determine the risk of lethargic driving. Ye et al. utilize Perspective-n-Point (PnP) to estimate the current head pose angle. Subsequently, fatigue is detected by assessing whether the head exceeds the angle of excessive deviation in head posture [24].

3.2.5. Facial expression. Some studies have shown that fatigued driving often leads to changes in the driver's facial expressions, which provide important clues for monitoring and recognising drowsiness. Common features of facial expressions include yawning closed eyes and so on. To realize the recognition of specific expression states, the classifier needs to be trained in advance for facial expression features and eye and mouth feature extraction. Facial expression classifiers often need to distinguish between multiple facial feature states. For example, eye feature extraction might detect blinking, opening, and closing eyes. In addition, facial expression extraction is the integration and comprehensive analysis of different facial feature information obtained from the driver's entire face. For example, the facial expression classifier can recognize that the driver has his eyes closed and his head is facing downward so that the driver is tired. By comprehensively analyzing facial expressions, the system can provide a more comprehensive assessment of the driver's fatigue state.

In 2015, Zhang et al. researched a deep learning-based model for recognising yawning in drivers. The proposed system integrates face recognition, nose detector, nose tracker, and yawn detector [25]. However, this system extracts features from a single image, and yawning is a continuous behaviour and not a static one. Since most existing methods for detecting yawning only use the spatial features of static images, more and more researchers have started to consider temporal deep models to better recognise drivers' yawning. In 2017, Zhang et al. used CNN to extract spatial image features and stacked LSTM layers in GoogLeNet to detect yawning [26]. Yang et al. proposed a 3D deep learning network with a low temporal sampling rate (3D-LTS) for facial motion recognition. It chooses a key image selection algorithm that quickly eliminates redundant images while achieving better performance [27]. Saurav et al. replaced LSTM with bidirectional LSTM (Bi-LSTM) [28].

3.2.6. Multi-feature detection. A single feature can affect the results of a model due to environmental changes or equipment. On the other hand, multi-feature fusion can use more comprehensive facial information, including eyes, mouth, expressions, expressions, head positions, and other different features, to improve the performance and robustness of the model. In addition, there are advantages to using the deep learning model in multi-feature fusion fatigue driving detection. Deep Learning can extract a variety of features from the driver's facial image. Then, the system integrates information from different facial features by using deep neural network structures such as feature cascades and fusion layers to ensure that different features work together effectively. Besides, since the contribution of different facial features may vary, a feature weighting mechanism needs to be introduced to optimize the feature fusion process. This can be achieved by training a weight distribution network or by manually adjusting the weight parameters. To sum up, using a variety of facial features to identify signs of fatigue, it is designed to improve the accuracy and reliability of the driver's state. The following are some studies of multi-feature fatigue detection using different depth learning algorithms.

Li and Bai proposed an algorithm based on MTCNN-PFLD-LSTM to recognise drowsy driving using deep learning techniques. The MTCNN is used to identify facial regions, while the PFLD model is used to identify eye, mouth, and head key points as well as spatial pose angles. These key points are then fed into LSTM for fatigue detection. By optimising the weights of the loss function, high recognition efficiency and accuracy can be achieved and can be applied to devices with less computational resources [29].

Huang et al. Introduced a multi-granularity Deep Convolutional Model based on feature Recalibration and Fusion for driver fatigue detection (RF-DCM), which contains three subnetworks: Multi-granularity Extraction sub-Net (MEN), Feature Recalibration sub-Net (FRN) and Feature Fusion sub-Net (FFN). The MEN and FRN participate in the feature extraction process, utilizing partial facial information and recalibrating local feature weights, respectively. On the other hand, the FFN operates in the subsequent steps of feature fusion, addressing challenges associated with integrating global and local features to achieve more comprehensive representations [30].

Dua et al. used four deep learning models, including AlexNet, VGG-FaceNet, FlowImageNet, and ResNet, to build a driver fatigue detection system. The four models deal with background and environmental changes, extract facial features, process behavioral features and head movements, and process gesture features, respectively [31].

Zhuang and Qi used a pseudo-3D convolutional neural network and attention mechanisms for fatigue driving detection. The P3D-CNN and P3D-Attention modules were used to study spatiotemporal features, and the feature correlation was improved through dual-channel attention and adaptive spatial attention [32].

Ghourabi et al. combined the three features of closed eyes, yawning, and multiple perception (MLP) and K-nearest (K-NN) classification techniques to detect fatigue states [33].

3.3. Summary and comparison of existing studies

In comparing research results, various studies employ different evaluation metrics, predominantly including Accuracy and F1-Score to represent performance. As given in Eq.1, True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) represent different outcomes of predictions. TP indicates that the correct prediction is a positive class, while TN indicates correct but negative. Conversely, FP indicates that the wrong prediction is a positive class, and FN indicates wrong but negative. The evaluation metrics mentioned are as follows:

Accuracy: The proportion of the model that is correctly predicted across all samples as in equation (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$(1)$$

F1-Score: The measure that takes into account both precision and recall as in equation (2, 3, 4).

$$Precision \; = \; \frac{TP}{TP \; + \; FP} \tag{2}$$

$$\mathrm{Re}call \; = \; \frac{TP}{TP \; + \; FN} \tag{3}$$

$$F1 - Score \; = \; \frac{2 \times Precision \; \times \; \mathrm{Re}call}{Precision \; + \; \mathrm{Re}call} \tag{4}$$

Table 1 summarizes the previously mentioned papers that use the deep learning method to derive fatigue-driving detection results through facial characteristics.

Table 1. Summary of previous studies about detecting fatigue driving based on facial features using deep learning

Reference	Authors	Methods	Drowsiness Measures	Performance
[21]	Quddus et al.	CNN, LSTM	Blinking, eye closure	Accuracy = 95.00% - 97.00%
[22]	Zhao et al	EM-CNN	PERCLOS, POM	Accuracy = 93.62%
[23]	Ping et al	MTCNN, Improved AlexNet based on the Ghost module	PERCLOS, PMOT	Accuracy = 93.50%
[29]	Li and Bai	MTCNN, PFLD, LSTM	Eye, Mouth, head, spatial attitude angles	Accuracy = 99.22%
[30]	Huang et al	RF-CNN	Eyes, mouth, glabella	F1-score = 89.42%
[31]	Dua et al	AlexNet, VGG-FaceNet, FlowImageNet, ResNet	Yawing, hand gestures and environment	Accuracy = 85.00%
[32]	Zhuang and Qi	Pseudo-3D CNN	Yawn, blinks, and head characteristic movements	F1-score = 99.89%
[33]	Ghourabi et al	MLP, KNN	Eye closure, yawning	Accuracy = 94.31% F1 - Score = 79.00%

4. Conclusion

Driver facial feature fatigue driving detection methods based on deep learning have made significant progress in improving accuracy and real-time performance. However, there are still some key challenges to further development in this area.

Abundant training data is important for fatigue detection, but it is not easy to obtain diverse, high-quality datasets. At present, public datasets mainly rely on simulated driving environments, which are very different from complex real-world driving environments. To ensure that the model can perform well in a variety of real-world situations, the dataset covers a variety of driving scenarios, including day, night, different weather conditions, etc. In addition, considering that the driver's age, gender, cultural background, and driving habits will have a certain degree of influence on the driving data. Therefore, the dataset should be continuously enriched and optimized to make the model more adaptable to different groups of people and different driving environments. Furthermore, fatigue detection requires high-quality data, because data quality directly affects the performance and generalization ability of deep learning models. The standard for high-quality data is high-resolution, low-noise, and covers different lighting and environmental conditions. In addition, there is a problem of subjectivity in the labeling of fatigue states in the dataset. Labeling of fatigue states often relies on the subjective assessment of the observer, which can lead to inconsistencies and ambiguity in the labeling.

Multi-feature fusion has obvious advantages in improving model performance and adaptability, but there are still potential challenges. Multi-feature fusion usually involves a more complex computational process and is more cumbersome. As a result, this is a challenge for both the researcher's expertise and the configuration of the environment.

In practical applications, the real-time nature of the algorithm is crucial. At present, most CNN-based feature extraction methods have good accuracy, but due to the deep network structure, the real-time

performance is poor. Second, it is necessary to optimize deep learning models further to ensure efficient performance in resource-constrained embedded systems.

References

[1]   Ansari S, Du H, Naghdy F, Stirling, D 2023 Factors Influencing Driver Behavior and Advances in Monitoring Methods. Lecture Notes in Intelligent Transportation and Infrastructure. Springer, Cham

[2]   Craye C, Rashwan A, Kamel M S, Karray F 2016 A multi-modal driver fatigue and distraction assessment system. International Journal of Intelligent Transportation Systems Research 14 173-194

[3]   Li X X 2021 Research on detection of Fatigue driving based on deep learning. University of Science and Technology of China

[4]   Li R, Chen Y J, Zhang L H 2021 A method for fatigue detection based on Driver's steering wheel grip. International Journal of Industrial Ergonomics 82

[5]   Zhu M 2021 Vehicle driver drowsiness detection method using wearable EEG based on convolution neural network. Neural computing & applications 33(20) 13965-13980

[6]   Yang Y, Li L, Lin H 2023 Review of Research on Fatigue Driving Detection Based on Driver Facial Features. Journal of Frontiers of Computer Science and Technology 17(6) 1249-1267

[7]   Ebrahim S 2023 A systematic review on detection and prediction of driver drowsiness. Transportation Research Interdisciplinary Perspectives 21 100864

[8]   Zhang R, Zhu T J, Zou Z L 2022 Review of research on driver fatigue driving detection methods. Computer Engineering and Applications 58(21) 53-66

[9]   Lecun Y, Bottou L, Bengio Y, Haffner P 1998 Gradient-based learning applied to document recognition. Proceedings of the IEEE 86(11) 2278-2324

[10]   Krizhevsky A, Sutskever I, Hinton G E 2012 ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 Curran Associates Inc p 1097-1105

[11]   Zeiler M D, Fergus R 2014 Visualizing and understanding Convolutional Networks. In ECCV LNCS 8689 p 818-833

[12]   Simonyan K, Zisserman A 2015 Very Deep Convolutional Networks for Large-Scale Image Recognition. In International Conference on Learning Representations.

[13]   Szegedy C, 2015 Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Boston, MA, USA p 1-9

[14]   He K, Zhang X, Ren S, Sun J 2016 Deep Residual Learning for Image Recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Las Vegas, NV, USA p 770-778

[15]   Hu Y, Luo D, Hua K, 2019 Overview on deep learning. CAAI Transactions on Intelligent Systems 14(01) 1-19

[16]   Van Houdt G, Mosquera C, Nápoles G 2020 A review on the long short-term memory model. Artificial Intelligence Review 53(8) 5929-5955

[17]   Fang H J, Dong H Z, Lin S X 2023 Driver fatigue state detection method based on multi-feature fusion. Journal of Zhejiang University (Engineering Science) 57(07) 1287-1296

[18]   Knipling Ronald R, Wierwille W W 1994 Vehicle-based drowsy driver detection: current status and future prospects IVHS America Fourth Annual Meeting.

[19]   Dinges D F, Grace R C 1998 Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Retrieved from https://api.semanticscholar.org/CorpusID:141471330

[20]   Ling Y C, Weng X X 2023 Efficient and robust driver fatigue detection framework based on the visual analysis of eye states. Promet-Traffic & Transportation 35(4) 567-582

[21]   Quddus A 2021 Using long short term memory and convolutional neural networks for driver drowsiness detection. Accident Analysis & Prevention 156 106107

[22] Zhao Z, Zhou N, Zhang L, Yan H, Xu Y, Zhang Z 2020 Driver Fatigue Detection Based on Convolutional Neural Networks Using EM-CNN. Computational Intelligence and Neuroscience 1-11

[23] Lou P, Yang X, Hu J W 2021 Fatigue driving detection method based on edge computing. Computer Engineering 47(07) 13-20+29

[24] Ye M, Zhan W, Cao P 2021 Driver fatigue detection based on residual channel attention network and head pose estimation. Applied Science 11(19) 9195

[25] Zhang W, Murphey Y L, Wang T, Xu Q 2015 Driver yawning detection based on deep convolutional neural learning and robust nose tracking. In 2015 International Joint Conference on Neural Networks (IJCNN) Killarney p 1-8

[26] Zhang W, Su J 2017 Driver yawning detection based on long short term memory networks. In 2017 IEEE Symposium Series on Computational Intelligence (SSCI) Honolulu, HI, USA p 1-5

[27] Yang H, Liu L, Min W, Yang X, Xiong X 2021 Driver yawning detection based on subtle facial action recognition. IEEE Transactions on Multimedia 23 572-583

[28] Saurav S, Mathur S, Sang I, Prasad S S, Singh S 2019 Yawn detection for driver's drowsiness prediction using bi-directional LSTM with CNN features. In International Conference on Intelligent Human Computer Interaction.

[29] Li X Bai C 2021 Research on driver fatigue driving detection method based on deep learning. Journal of the China Railway Society 43(06) 78-87

[30] Huang R, Wang Y, Li Z, Lei Z, Xu Y 2022 RF-DCM: Multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection. IEEE Transactions on Intelligent Transportation Systems 23(1) 630-640

[31] Dua M, Shakshi, Singla R, Raj S, Jangra A 2020 Deep CNN models-based ensemble approach to driver drowsiness detection. Neural Computing and Applications 33 3155-3168

[32] Zhuang Y, Qi Y 2021 Driving fatigue detection based on pseudo 3D convolutional neural network and attention mechanisms. Journal of Image and Graphics 26(1) 143-153

[33] Ghourabi A, Ghazouani H, Barhoumi W 2020 Driver drowsiness detection based on joint monitoring of yawning, blinking, and nodding. In 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP) Cluj-Napoca, Romania p 407-414