Engineering Applications of Artificial Intelligence 116 (2022) 105399

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Survey paper

A review of driver fatigue detection and its advances on the use of RGB-D camera and deep learning

Fan Liu [a,b,c,*], Delong Chen [b], Jun Zhou [d], Feng Xu [b]

[a] Key Laboratory of Water Big Data Technology of Ministry of Water Resources, Hohai University, Nanjing, China
[b] College of Computer and Information, Hohai University, Nanjing, China
[c] Science and Technology on Underwater Vehicle Technology Laboratory, Harbin, China
[d] School of Information and Communication Technology, Griffith University, Nathan, Australia

ARTICLE INFO

Keywords:
Intelligent transportation system
Driver fatigue detection
Abnormal behavior detection
Information fusion
RGB-D
Deep learning
Computer vision
Convolutional neural network

ABSTRACT

Driver fatigue is an essential reason for traffic accidents, which poses a severe threat to people's lives and property. In this review, we summarize the latest research findings and analyze the developmental trends of driver fatigue detection. Firstly, we analyze and discuss four types of different fatigue detection technologies based on driver physiological signals, behavior features, vehicle running features, and information fusion, respectively. Then, we focus on RGB-D camera and deep learning which are two state-of-the-art solutions in this field. Finally, we present the work on integration of RGB-D camera and deep learning, where Generative Adversarial Networks and multi-channel schemes are utilized to enhance the performance. We conducted experiments to show that the fatigue features extracted by Convolutional Neural Networks are superior to traditional handcrafted ones while single features cannot guarantee robustness. Moreover, the latent fatigue features extracted by deep learning methods have been demonstrated to be effective for fatigue detection.

1. Introduction

Nowadays, people are more and more sleep-deprived due to hectic lifestyle in proliferating economies. With increasing amount of vehicles on the road, driver fatigue becomes a significant cause of accidents. According to National Highway Traffic Safety Administration (NHTSA) (Asleep, 2020), driver fatigue has been one of the major threats to life safety and the economy. The American Automobile Association also noted that 21% of fatal crashes result from driver fatigue (Tefft, 2012). Therefore, driver fatigue detection has become a vital task for preventing traffic accidents.

In the past several decades, numerous fatigue detection methods and technologies have been developed. These methods have been covered by survey papers from different perspectives, for example, psychology and causing of driver fatigue (Lal and Craig, 2001a), measurements of physiological signals (Sanjaya et al., 2016), fatigue detection based on driver's behavior or performance (Wang et al., 2006), and a combination of all these perspectives (Shi et al., 2017; Stork et al., 2015). Sahayadhas et al. (2012) reviewed and discussed the sensors used by different measures for fatigue detection. Some surveys (Alsibai and Manap, 2016; Williamson and Chamberlain, 2005; Mashko, 2015; Chacon-Murguia and Prieto-Resendiz, 2015; Sikander and Anwar, 2018) covered representative systems, devices, tools, applications, and problems in driver fatigue detection. Others (Owen et al., 2015; Meng et al., 2015; Golz et al., 2010) focus on special topics for professional drivers such as truck, taxi, and racing drivers. Golz et al. (2010) reviewed and evaluated some commercial devices to meet the requirements of the mining industry. In Koesdwiady et al. (2017b), Koesdwiady et al. analyzed recent trends in driver safety monitoring systems and reviewed some driver fatigue detection techniques. More recent works on the advantages and disadvantages of features, classifiers, accuracy, system parameters, and environment can be found in Kaplan et al. (2015), Ramzan et al. (2019), Němcová et al. (2021) and Wang et al. (2021).

In recent years, fatigue detection has entered a new development period due to wide use of RGB-D camera and deep learning technologies. However, all above-mentioned survey papers have not reviewed or discussed the role of these two technologies for fatigue detection. This review fills the gap by not only analyzing classic driver fatigue detection methods, but also presenting the trends brought by RGB-D camera and deep learning in the field. Furthermore, combination of these two technologies is evaluated in experiments.

As shown in Fig. 1, we classify driver fatigue detection methods into direct and indirect categories according to their relationship with fatigue. The direct methods are further divided into the physiological signal-based and driver behavior features-based. The indirect methods mainly refer to those using vehicle driving features which cannot

* Corresponding author at: College of Computer and Information, Hohai University, Nanjing, China.
E-mail address: fanliu@hhu.edu.cn (F. Liu).

https://doi.org/10.1016/j.engappai.2022.105399
Received 5 March 2022; Received in revised form 29 June 2022; Accepted 26 August 2022
Available online xxxx

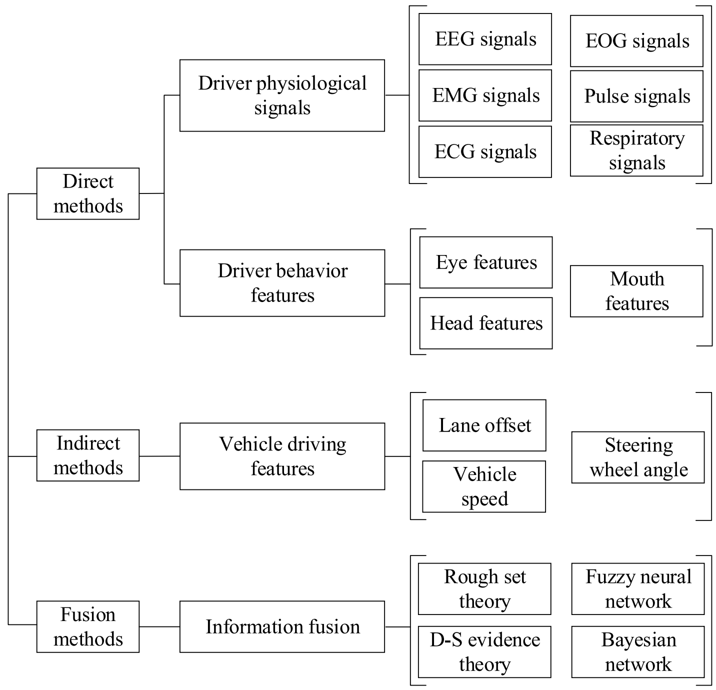

Fig. 1. Different categories of fatigue detection methods.

directly reflect the degree of fatigue. Both direct and indirect methods focus on single fatigue characteristics, which sometimes cannot guarantee robustness. This drawback can be addressed by information fusion-based methods which integrate multiple fatigue features to enhance reliability of detection.

This paper is an extension of our previous work (Liu et al., 2019a), which is enriched by reviewing more recently published papers on driver fatigue detection. Moreover, we add comparison and analysis of different fatigue detection technologies, especially for methods based on RGB-D camera and deep learning technologies.

Leveraging the analysis of research trends, we propose a novel driver fatigue detection framework by integrating RGB-D camera and deep learning. Additionally, we construct a driver fatigue dataset and conduct experiments to compare multiple deep learning solutions for driver fatigue detection, such as Deep Convolutional Generative Adversarial Networks (DCGAN) (Radford et al., 2016), Deep Belief Net (DBN) (Hinton et al., 2006), Stacked Denoising Auto-encoders (SDAE) (Vincent et al., 2010), and CNN (Krizhevsky et al., 2012). Experimental results show that the deep fatigue features are superior to traditional visual fatigue features. Moreover, data augmentation based on multi-channel scheme and DCGAN (Radford et al., 2016) are proved to be effective. Then, the information fusion strategy can further enhance the robustness of fatigue detection.

The main contributions of this paper are as follows:

1. We overview and analyze four different types of fatigue detection technologies.
2. We review and discuss RGB-D camera and deep learning technologies as two state-of-the-art solutions for driver fatigue detection.
3. We present the idea of using RGB-D camera and deep learning technologies simultaneously and point out it is a valuable future research direction.
4. Moreover, we validate and analyze the feasibility of multiple deep learning solutions for driver fatigue detection.

The rest of the paper is organized as follows. In Section 2, we present the classic methods and products in driver fatigue detection. Then we analyze the advantages and disadvantages of different fatigue detection technologies in Section 3. Section 4 introduces new developing trends and Section 5 validates the proposed solutions by experiments. Finally, the conclusions of our work are drawn in Section 6.

2. Classic methods and products

As shown in Fig. 1, fatigue detection methods can be divided into three categories: direct methods, indirect methods, and fusion methods. Direct methods are based on driver physiological signals and driver behavior features that are directly related to driver's fatigue state. The indirect methods refer to those based on vehicle driving features which can indirectly reflect fatigue state. Both direct and indirect methods rely on a single fatigue feature. On the contrary, the fusion methods can integrate multiple fatigue features for the final prediction. This section describes the classic fatigue detection methods based on physiological signals, driver behavior features, vehicle running status and information fusion respectively.

2.1. Fatigue detection based on physiological signals

Studies show that the physiological signal indicators deviate from the normal values during fatigue (Shen et al., 2008). Therefore, whether drivers are in a fatigue state can be determined by the changes of the physiological signals. As the main bio-electric signals, the driver's electromyogram (EMG), electrooculogram (EOG), electrocardiogram (ECG), and electroencephalogram (EEG) have already been leveraged in fatigue detection. These methods are regarded as the fatigue detection based on physiological signals.

The amplitude changes and frequency variation of EMG signals can reflect the fatigue state, i.e. the amplitude increases with decreased frequency during fatigue. The research by Naeije and Zorn (1982) proved that significant changes in the EMG power spectrum may occur during muscle fatigue. Based on this study, Hostens and Ramon (2005) proposed an evoked potential method that can detect driver's fatigue state during long-distance driving. Petrofsky and Lind (1980) studied the impact of temperature on the frequency and amplitude of EMG power spectrum. Recently, Lu et al. (2021) proposed to embed surface electromyogram (sEMG) sensors on a steering wheel. This method significantly improved the convenience and user experience, but brought additional noise in the collected sEMG signal. Besides EMG, the tendency of eye closing during fatigue can also be effectively detected by EOG signal. For example, eye blink can be classified based on peak amplitude, rise time, and fall time of the EOG waveform (Ohsuga et al., 2007; Noguchi et al., 2007).

Since there were significant differences between the status of the awakeness and tiredness for ECG signals (Calcagnini et al., 1994), some representative characteristics of ECG signals such as low-frequency energy (LF), very low-frequency energy (VLF), high-frequency energy (HF), and LF/HF ratio have been used to detect fatigue. Rogado et al. (2009) proposed to place the ECG electrode on the steering wheel grip to obtain the real-time ECG data and heart rate variability (HRV). Jeong et al. (2007) presented a driver stress index provision system with a new ECG measuring method, which converts the ECG signals to HRV for analyzing the time and frequency domain. Buendia et al. (2019) also demonstrated the effectiveness of HRV by finding the relationship between HRV and sleepiness.

Lal and Craig (2001b) and Lal et al. (2003) found the correlation between the driver's fatigue state and the EEG. Experimental comparison and analysis show that the δ and θ waveforms in the EEG hardly change when the driver is in the early stage of fatigue. But when the driver is extremely tired, the waveform in the EEG becomes greatly deepened. According to the study of full-night spectroscopy (Cajochen et al., 1999), topographical and temporal impact can be found on EEG activity during sleep when the driver is lack of sleep. Therefore, EEG signal has been widely used for evaluating driver fatigue state. Luo et al. (2019) proposed to classify the fatigue state by adaptively extracting multiscale entropy features of the forehead EEG data. Zuraida et al. (2019) tried to identify the best parameters for EEG based fatigue detection methods. In their experiments, EEG changes related to fatigue were only found at night. Further study from the perspective of complex network theory reveals the relationship between brain network and the fatigue behavior (Han et al., 2019b; Cai et al., 2019). Feature extraction network and feature selection network can be used to extract more discriminative features from EEG signals (Tuncer et al., 2020). Different from the widely used spectral analysis of EEG signals, Tran et al. (2008) proposed a sampling entropy and a second-order difference plot method, which are quantized by Central Trend Measurement (CTM) and used to nonlinearly analyze and process EEG signals. In addition, according the finding of Kong et al. (2017), the delta and alpha bands of EEG synchronizations in frontal and parietal lobe significantly increase with fatigue. They used Mean Phase Coherence (MPC)-based inter/intra-region phase synchronization and functional units (FUs) to detect fatigue effectively.

For industry products, the Smart Car developed by MIT measures the driver's ECG signals, respiratory speed, skin resistance, and other factors to estimate the driver's fatigue state (Healey and Picard, 2000). The Japanese Pioneer Corporation invented a product by detecting the speed of a driver's heartbeat to prevent drowsiness (Pioneer Corporation, 2020). These products directly put sensors on driver's body, leading to unpleasant user experience and sometimes even disturbing the driving.

2.2. Fatigue detection based on driver's behavior features

Driver's fatigue state can also be detected based on the driver's behavior features, for example, changes in the driver's eyes, mouth, head, and so on. Fig. 2 shows several Driver Danger Monitors from American Attention Company. Early version (DD850) captures information about driver's eyes at night from an infrared camera. Upgraded products such as MR688 and gogo850 (Nanjing Yuanqu Technology Co., Ltd., 2020) work during the day as well. In the following, we will introduce the technical details of this type of methods.

The state of eyes is regarded as an important feature for fatigue detection. Wierwille et al. (1994) proposed PERCLOS as an important parameter for driver fatigue detection. PERCLOS is the percentage of duration of closed-eye state in a specific time interval (1 min or 30 s) (PERCLOS, 1998). The PERCLOS-based method has been recognized as the most effective on-board real-time driver fatigue assessment method. EM, P70, and P80 are commonly used PERCLOS criteria which indicates that eye closure is more than 50%, 70% and 80% respectively. U.S. Federal Highway Administration (FHWA) and U.S. National Highway Traffic Safety Administration (NHTSA) compared nine fatigue detection indexes and found P80 showed the greatest correlation with driver fatigue. Pauly and Sankar (2015) adopted PERCLOS for fatigue detection, in which they used a Haar based cascaded classifier and an HOG based SVM (Chang and Lin, 2011) classifier for eye rotation tracking and eye-blink detection respectively, then calculate PERCLOS to predict the driver fatigue state.

In addition to PERCLOS, other important eye features for fatigue detection include eye gaze direction, blink frequency, and pupil characteristics. Choi and Kim (2014) identify driver's pupil position and determine the fatigue state by the changes of driver's sight-line. Since the blink frequency under fatigue state is significantly higher than normal, Suzuki et al. (2006) proposed a multiple regression model to detect the fatigue by simultaneously calculating blink frequency and PERCLOS. After evaluating 18 different eye features, Friedrichs and Yang (2010) found that eye opening speed features and PERCLOS achieved best fatigue detection accuracy.

Although the eye feature is highly useful, it is often affected by reflections of glasses or the occlusion of sunglasses. The relatively small eye region also brings difficulty for analyzing eye state. In contrast, other driver's behavior features such as head or mouth state are easier to capture. Since driver's head position is different in fatigue and awake state, driver's fatigue can be detected based on the nod frequency during a specific period of time, e.g. using a fuzzy classifier (Bergasa et al., 2006). Driver's head position and angle of rotation were also used to detect driver's attention state (Murphy-Chutorian and Trivedi, 2010). For mouth feature, driver's specific mouth movements, such as yawning, may reflect the fatigue state. So far both geometric and texture mouth features have been extracted for yawning detection (Rongben et al., 2004; Fan et al., 2007). Moreover, due to the temperature difference inside and outside the mouth, yawning can also be detected from thermal images (Knapik and Cyganek, 2019).

2.3. Fatigue detection based on vehicle driving features

Vehicle's driving state can somehow reflect driver's fatigue. Therefore, driving features such as lane offset, steering wheel angle, and vehicle speed have been employed to detect fatigue. Fig. 3 shows a few representative products, such as SafeTRAC (SafeTrac, 2020) from American AssitWare Company, and the ADAS-P9 (ADAS, 2020) from Israel Mobileye Company (Mobileye Company, 2020). These products perform fatigue detection based on the offset state of the running vehicle wheel.

Line offset measures the deviation of vehicle from the existing track. The vehicle's driving path is mainly monitored by computer vision technology. Dingus et al. (1998) modeled the correlation between the PERCLOS and driver's lane offset produced by fatigue and proposed that line offset information can be utilized as an indicator to detect driver fatigue. Based on this idea, Chang et al. (2008) applied a Radial Basis Probability Network (RBPN) to recognize the lane offset. Matsushita and Miura (2011) developed a particle filter based lane detection method. It treats the extracted road edge information as several particles and uses particle filtering to detect their changes.

When the steering wheel has little variation for a while, and abruptly changes within a vast range, it is likely to be attributed to the fatigue state. According to this assumption, Takei and Furukawa (2005) proposed a fatigue detection approach based on steering wheel angle, which pre-processed the angle signals by Fast Fourier Transforms (FFT) and Wavelet Transforms and then extracted relevant features. By monitoring the vehicle's trajectory and steering wheel rotation angle, Zhong

Fig. 2. Different fatigue detection products based on driver behavior features (DD850, gogo850, and MR688).
Source: Images extracted from Nanjing Yuanqu Technology Co., Ltd. (2020).

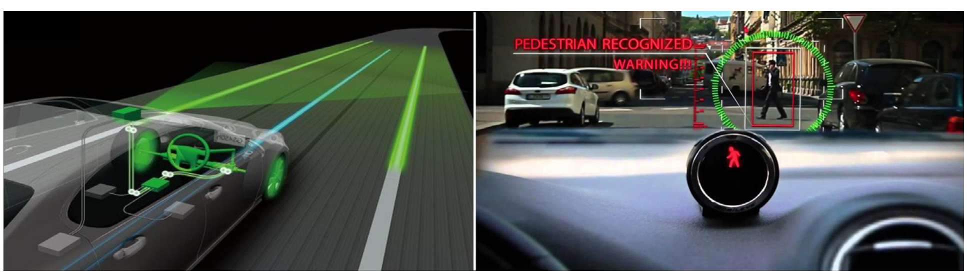

Fig. 3. Advanced Driving Assistant System (ADAS-P9). Image extracted from ADAS (2020).

et al. (2007) employed energy analysis and wavelet analysis techniques to evaluate whether the driver was fatigued or not. Eskandarian and Mortazavi (2007) carried out experiments of drowsiness detection for truck drivers based on artificial neural networks, which showed that steering signal difference could be effectively used for fatigue detection.

During fatigue state, the driver cannot estimate the speed accurately. Therefore, the driving speed can also be used as a feature to determine whether the driver is fatigued (Sandberg and Wahde, 2008). However, it is difficult to accurately infer the fatigue state based on the information of vehicle speed solely. This is because the vehicle speed is highly susceptible to external factors, such as traffic conditions and personal driving styles. One solution is introduced by Hu et al. in Hu et al. (2017), where they used real-world testing data to establish personalized driving models. But this model only uses speed, throttle position and brake pressure as input, which greatly limits the model ability. Therefore, researchers began to combine these features to provide more reliable detection results.

2.4. Fatigue detection based on information fusion

Fatigue is a physiological phenomenon with individualized characteristics, i.e. different drivers have different fatigue features. Therefore, single fatigue feature-based methods may not be suitable for all drivers. In addition, detection methods based on single features are easily affected by the interference of different environment conditions. Therefore, it is hard to guarantee the validity and credibility of the results.

In recent years, information fusion technologies have played an increasing important role in driver fatigue detection. Information collected by different sensors produce different fatigue features, which can be fused for fatigue detection to make up the shortcomings of single feature and improve the robustness of fatigue detection model. For example, Seeing Machines exploited the faceLAB system (Fig. 4) (FaceLab, 2020) by integrating the features of PERCLOS, sight-lines and blink frequency. Similarly, the AWAKE (System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation) project (TRIMIS, 2020) implemented an effective driver fatigue detection system by combining the features of eyes, steering wheel angle, steering wheel force and lane offset.

Information fusion technologies for fatigue detection can be classified into rough set theory, dynamic Bayesian network, D-S evidence theory, fuzzy neural network, and other fusion technologies. Du et al. (2011) utilized a kernelized fuzzy rough sets to evaluate the quality of candidate features and select the useful subset for fatigue detection system. In recent years, rough set theory has been exploited to filter different fatigue signals (Ye and Zhao, 2018) or compute their weights (Chen et al., 2018) to further enhance the robustness of driver fatigue detection.

Yang et al. (2005) proposed D-S evidence theory to combine different features related to driver fatigue, such as driver's sleep quality, body temperature, mental state and so on. It was pointed out that

Fig. 4. Driver fatigue monitor (faceLAB).
Source: Image extracted from FaceLab (2020).

Table 1
Comparison of different fatigue detection methods.

Detection technology	Contact	Accuracy	Practicability
Driver physiological signals	✓	Very high	Low
Driver behavior features	×	High	Very high
Vehicle driving features	×	Medium	High
Information fusion	×	Very high	Very high

Table 2
Fatigue detection accuracies using different input signals (Ahn et al., 2016).

Input signal	Classification accuracy
EEG	59.7
ECG	64.5
fNIRS	66.8
EEG + ECG	69.0
EEG + fNIRS	68.3
ECG + fNIRS	69.2
EEG + ECG + fNIRS	**75.9**

the D-S evidence theory not only effectively integrates lane offset and eye features, but also successfully handles noisy data (Li et al., 2014). The D-S evidence theory can also be integrated with Fuzzy Neural Network (FNN). For example, Zhu et al. (2017) extracted multiple facial features and used the FNN with adaptive and self-learning capabilities to acquire a basic probability distribution of evidence. Then, a discount strategy (Shafer, 1976) were exploited to amend the raw evidence and detect the driver's fatigued state.

One common limitation of the above information fusion methods is they only explore the relationship between different fatigue features, but do not consider their time dependencies. To overcome such limitation, Yang et al. (2010) proposed to use a dynamic Bayesian network to achieve fatigue driving detection with PERCLOS, EEG, ECG, sleep quality (SQ), circadian rhythm (CR), working environment (WE) and eye movement (EM) as input. Among these factors, EM, ECG, and EEG can be obtained directly; SQ, WE and CR contain important context information; fatigue and alarm states are used as the hidden variables. Similarly, dynamic Bayesian network can be used with sleep quality, the road environment, and the driving duration as input (Bani et al., 2019) or with other variables related to vehicles, drivers and environment (Chhabra et al., 2019; Al-Sultan et al., 2013).

3. Comparison of different fatigue detection technologies

In Section 2, we have reviewed four types of existing approaches for driver fatigue detection. Technical details of different approaches are presented. Now, we turn to present an in-depth analysis of the characteristics of these approaches. First, the pros and cons of the four types of approaches are compared. We summarize their properties from different perspectives in Table 1, where "Contact" means whether the sensor directly contacts the driver. We note that since mainstream of physiological signal-based approaches employs contact devices, physiological signal-based approaches are marked as "contact" in Table 1. "Accuracy" indicates the accuracy of fatigue detection system, and "Practicability" is the system adaptation capability. In the following, we compare and discuss four types of fatigue detection methods in detail, and give corresponding results and analysis.

3.1. Fatigue detection based on driver physiological signals

In recent years, thoughtful experiments have been conducted to compare the effectiveness of different physiological signals for driver fatigue detection. For example, Ahn et al. (2016) compared EEG, ECG, and functional near infra-red spectroscopy (fNIRS) signals. As is Table 2, their experimental results on eleven subjects shows that, among the most popular two kind of physiological signals, ECG outperforms EEG by nearly 5%. They also find that fNIRS signal can yield better results. Later, Du et al. (2017) experimentally compared the result of EOG and EEG signals, which demonstrated that EOG is superior to EEG.

Though the EOG-based fatigue detection approaches have shown advantages in terms of accuracy, the complicated instruments may affect the driving due to the visual interference. Compared to the EOG-based methods, the ECG based techniques are easier to carry and operate. However, the ECG signals have lower sensitivity, and the discrepancies of different drivers may produce a significant error in the driver's heart rate (Naaz and Singh, 2014). Pulse detection is simpler and more sensitive, but it is also uncomfortable for the driver because the sensor needs to be in direct contact with the driver. Furthermore, the individual difference of driver's pulses leads to inaccuracy. The fatigue detection methods based on the EMG signals ensure objective authenticity (Raez et al., 2006), but the detection of the EMG signals invades the driver's skin and may have a negative influence on the driver's driving safety.

In general, fatigue detection by physiological signals are more accurate than other non-contact approaches. Nonetheless, acquiring physiological signals requires electrodes to contact the driver's body, which leads to some interference and additional burdens for driving. Therefore, the practicality of this kind of methods is poor. Furthermore, though combining multiple physiological signals can lead to superior performance (Table 2), too much attached sensors could make the driver very uncomfortable. Based on this fact, more and more researchers are turning to exploiting human behavioral characteristics and vehicle driving characteristics for fatigue detection since they can be extracted by using non-contact devices. Some researchers tried to extract physiological signals by a non-contact camera. For example, Tsai et al. (2020) proposed a vision-based remote photoplethysmography (rPPG) signal measurement system for driver fatigue detection. However, the robustness of this type of method is poor.

3.2. Fatigue detection based on driver's behavior features

To compare the effectiveness of different driver behavior features, Qiao et al. (2016) conducted experiments using 500 blinking, 500 head shaking, and 500 yawning to detect drier fatigue. Table 3 listed the

Table 3
Test results of driver's behavior features (Qiao et al., 2016).

Behavior feature	Eye state	Head state	Mouth state
Detected fatigue	451/500	469/500	438/500
Accuracy	90.2%	**93.8%**	87.6%

detection accuracies, which show that the eye state and head state are more sensitive than mouth state. Generally, the eye features are mostly adopted for fatigue detection among existing literature due to its high accuracy and sensitivity (Gu et al., 2002). However, the acquisition of eye features is susceptible to factors such as sunlight, obstacles, and whether the driver wears glasses. The characteristics of the driver's head are also useful for fatigue detection, but it is less expandable and sensitive. Therefore, these features are often used as an assistance to fatigue detection. As the mouth characteristics are susceptible to the external environment, the fatigue detection methods based on mouth features are also regarded as an auxiliary way for fatigue detection.

In short, although computer vision-based techniques have been successfully employed to determine the fatigue state by monitoring the driver's eyes, head, and mouth features in a real-time and non-contact way, they can be easily influenced by different driving circumstances and environments, causing reduced performance (Abbas and Alsheddy, 2021).

3.3. Fatigue detection based on vehicle driving features

As far as the vehicle driving features, Sandberg and Wahde (2008) tested four leading drowsiness indicators of vehicle driving features, i.e., vehicle speed, lateral position, steering wheel angle, and yaw angle of the vehicle. They found that the lateral position could achieve relatively good fatigue detection results. It can be noticed that most driving behavior based fatigue indicator is context-dependent to some degree. To demonstrate this point, we collect 365 daytime road images and 570 night road images and use Hough transform (Illingworth and Kittler, 1988) to perform lane detection. The lane line detection results are shown in Fig. 5. The accuracies of lane detection at night drops 1.46% (from 93.76% to 92.3%) compared to that of daytime. Furthermore, when the vehicle is out of control, the result of fatigue driving detection is meaningless. Although the angle data of the steering wheel and vehicle speed collected by the sensor have better sensitiveness without influencing normal driving, they are easily affected by external factors such as different driving habits, the driving environments, and the type of vehicle. More importantly, it is hard to judge whether the driver is fatigued while driving at a low speed. For all the above reasons, fatigue detection based on vehicle speed (Sandberg and Wahde, 2008) has the lowest accuracy.

In a broad sense, the most significant advantage of the fatigue detection methods based on the driver's behavioral features and vehicle driving features is that the data acquisition is easy. However, the difficulty of these methods lies in how to divide the driver's state into "fatigue" and "non-fatigue", under different driving habits, vehicle variety, and road conditions. Therefore, it is necessary to design different criteria and set different "thresholds" for these methods.

3.4. Fatigue detection based on information fusion

The most important problem of data acquisition during fatigue detection is uncertainty. The rough set theory (Du et al., 2011) can be specifically designed to cope with this uncertainty at a low cost. Although the uncertainty can be analyzed objectively, it does not consider the expert experience. Dynamic Bayesian network can cope with the sample sets with incomplete or lost data. Therefore, the driver fatigue detection model using a dynamic Bayesian network (Knapik and Cyganek, 2019; Dingus et al., 1998) can effectively promote the reliability and robustness of early warning. The occurrence of fatigue is a gradual process, which requires training the network model every once in a while. Otherwise, it will seriously affect the reliability of detection. The D-S evidence theory (Yang et al., 2005) is regarded to have robust decision processing ability, which can efficiently combine diverse information from many kinds of sensors by uncertainty reasoning. Nonetheless, it still cannot get rid of the limitation of strict combination conditions and exponential computation increase. By combining the benefits of fuzzy theory and neural networks, FNN can effectually exploit the correlative experts' experience to detect the fatigue state. Moreover, it has high recognition accuracy and fast learning speed. The effectiveness of FNN is widely demonstrated. For example, in Dong et al. (2008), the FNN is used to fuse the curvature of up-eyelid and the distance of eyelid. Their results show that, the correct rate increases by almost 8% after FNN fusion, and the wrong rate decreases by nearly 11%. However, FNN also has many shortcomings, such as long training time, poor additivity, and complicated structure.

In general, traditional fatigue detection methods are limited with low fault tolerance when using a single source of information as the input. Instead, those information fusion based fatigue detection methods do not rely on any unique fatigue features, but perform comprehensive analysis of all the features of the driver to detect the fatigue state. Therefore, they have better applicability and fault tolerance ability.

4. Development trends

At present, with the rapid development of various new technologies, more fatigue detection methods emerged, which mainly reflects in the following trends.

4.1. RGB-D camera and low-cost solutions

Non-contact fatigue detection approaches are mostly based on driver behavior features. The drivers easily accept these approaches since they are convenient and only need to put the camera in front of drivers. However, most non-contact methods are based on conventional RGB camera, which acquire RGB images by projecting a three-dimensional scene onto a two-dimensional plane. As a result, relative angle and distance to the camera are lost in the imaging process. Moreover, driver's clothes and lighting variation produce diversified appearance information, which affects the extraction of the driver's fatigue features. Another problem of the RGB camera is how to obtain a striking image at night. Recently, aiming at the shortcomings of RGB camera, RGB-D camera have been increasingly applied in fatigue detection. For example, low-cost infrared depth camera Kinect from Microsoft has been used for fatigue detection (Marinello et al., 2014). In Fig. 6, we show some examples of RGB image, depth image, near-infrared image, and skeleton captured by Kinect.

The depth data not only effectively compensates for RGB image but also avoids confusion caused by the driver's cloths and lighting variations. Leveraging these strengths, many depth camera-based driver fatigue detection approaches have emerged in recent years. Since head pose can be effectively estimated by RGB-D camera, an efficient method is to first estimate the head pose using RGB-D images, then use this information to learn to predict driver fatigue (Cao and Lu, 2013; Wongphanngam and Pumrin, 2016; García et al., 2014). Head pose can also be integrated with other features for example, eye state (Zhang et al., 2015a), or a combination of eye, hand, head, and face features (Craye and Karray, 2015). Recently, Du et al. (2021) employed an RGB-D camera to extract heart rate, eye openness level, and mouth openness level, then used a recurrent neural network to fuse these multi-modal features. With the help of the depth images obtained from the RGB-D camera, the plane distance to the camera can be calculated to partially make up for the information missing from RGB cameras. Moreover, the depth data will not be interfered with by the appearance information. Therefore, more and more researchers focus on the utilization of RGB-D

Fig. 5. Lane detection at daytime and night. First row: input images, second row: detection results.

Fig. 6. RGB image, depth image, near-infrared image, and skeleton captured by Kinect.

camera in fatigue detection, which has become a novel and significant research direction.

Recently, some other low-cost solutions have also been adopted for fatigue detection, such as low-cost sensors, embedded devices and smartphones (Meireles and Dantas, 2019; Isaza et al., 2019; Shin et al., 2019; Xie et al., 2019). Sikander and Anwar (2021) proposed 3-D representation-based fatigue detection by utilizing fatigue-related facial action units, and exhibited its superiority to 2D representation. Inspired by these low-cost solutions, we also tried to design a Raspberry Pi based low-cost platform, which is shown in Fig. 7. An additional Movidius neural compute stick was plugged in the USB port to enhance the computing ability. Moreover, six infrared lamps were put around the camera to obtain better videos at night. As shown in Fig. 8, this kind of low-cost solution can also meet the fatigue detection requirement.

4.2. Deep learning

In recent years, many kinds of deep learning technologies have been developed. Deep learning stimulates the mechanism of the human brain to interpret the data and effectively extracts the characteristic of the data. Deep models automatically learn and establish fatigue detection standards from the training samples. By using powerful deep models, we can get rid of the dependence on those handcrafted fatigue detection standards. Moreover, deep learning technologies can learn the fatigue characteristics and standards for specific groups, which improves the adaption to different groups. Therefore, many motor corporations and research institutions have begun to exploit deep learning technologies to study driver fatigue detection. According to the aforementioned classification rule, deep learning-based fatigue detection approaches can also be divided into four types.

4.2.1. Driver physiological signals

Pioneer work of deploying deep learning models to extract features from EEG signals can be dated back to 2010, when Deep Belief Network (DBN) and Convolutional Neural Network (CNN) (Wulsin et al., 2011; Cecotti and Graser, 2010; Li et al., 2021) were used to extract features from EEG signals. DBN and CNN are usually tested on 2D image data, but they also showed strong ability to extract features from EEG

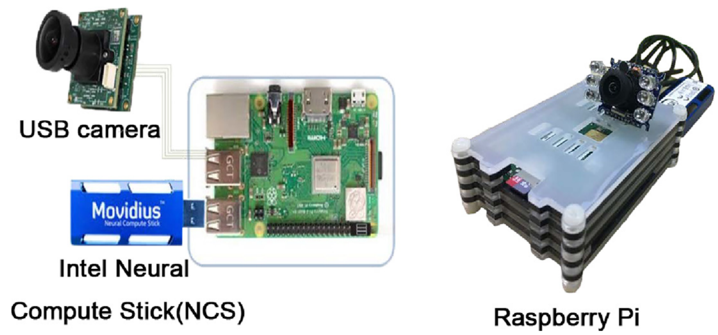

Fig. 7. A Raspberry Pi based low-cost platform.

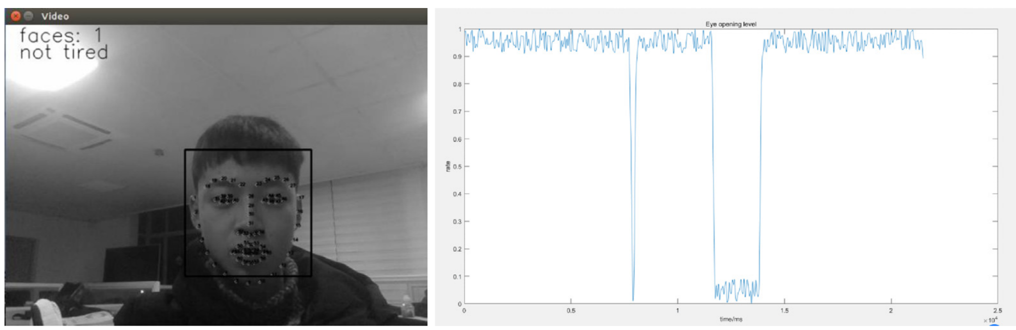

Fig. 8. Fatigue detection results based on Raspberry Pi. Left: detected face bounding box and landmarks. Right: estimated eye opening level.

Table 4
EEG signals based deep learning methods.

Ref.	Representation	Classifier	Accuracy
San et al. (2016)	Power Spectral density DBN	SVM	61.00% 73.29%
Hajinoroozi et al. (2015)	PCA DBN	Bagging	AZ score = 70.98% AZ score = 81.36%
Wen and Zhang (2018)	AutoEncoder	AdaBoost	95.00%
Wu et al. (2019)	AutoEncoder	Softmax	91.67%
Zhu et al. (2014)	AutoEncoder	Linear Regressor	Correlation Coefficient = 0.73
Rundo et al. (2019)	Stacked AutoEncoder	Softmax	100%
Ma et al. (2019b)	Modified-PCANet	SVM	95.14%
Panwar et al. (2019)	GAN	Softmax	AUC = 66.49%

signals. To deal with the driver fatigue detection task, a straightforward solution is to train an unsupervised deep model to extract features and feed the deep features into traditional classifiers. So far, deep features extracted by DBN have been used with SVM and Bagging classifier (San et al., 2016) and Hajinoroozi et al. (2015). Similarly, other unsupervised deep models such as AutoEncoder (AE) (Wen and Zhang, 2018; Wu et al., 2019; Zhu et al., 2014), Stacked AutoEncoder (SAE) (Rundo et al., 2019), PCANet (Ma et al., 2019b), and Generative Adversarial Network (GAN) (Panwar et al., 2019) were also introduced to extract deep features. We summarize the existing physiological signal-based deep driver fatigue detection approaches in Table 4.

As another emerging research trend, end-to-end deep learning models gradually began to be deployed. These methods use different signal processing methods such as covariance matrices (Hajinoroozi et al., 2017), Continuous Wavelet Transform (CWT) (Poorna et al., 2021), Differential Entropy (DE), and Power Spectral Density (PSD) (Ko et al., 2020; Zhang et al., 2020), and combine the advantage of both effective signal processing algorithms and powerful deep models. Zhang et al. (2020) combined DE and PSD with a Graph Convolutional Neural Network (GCNN) to extract features from spatial and temporal domains simultaneously. Similarly, an EEG-Based Spatio-Temporal Convolutional Neural Network (ESTCNN) was proposed by Gao et al. (2019). Furthermore, fussy neural network was combined with RNN and CNN to extract spatio-temporal features (Rundo et al., 2019; Ma et al., 2019b). In summary, some physiological signal-based deep methods mainly focus on learning temporal features (San et al., 2016; Hajinoroozi et al., 2015; Wen and Zhang, 2018; Wu et al., 2019; Zhu et al., 2014; Rundo et al., 2019; Ma et al., 2019b; Panwar et al., 2019; Hajinoroozi et al., 2017; Poorna et al., 2021; Ko et al., 2020), while some others paid attention to modeling spatial information such as dependencies between electrode nodes (Zhang et al., 2020; Gao et al., 2019; Liu et al., 2015; Du et al., 2020).

4.2.2. Driver behavior features

Eye state feature. Traditional eye state classification methods often suffer from illumination, occlusion, and pose variation in real-world driving conditions. In contrast, deep models are more robust and accordingly can produce higher accuracy. Kim et al. (2017) first attempted to utilize CNN to perform eye state classification. Later, Zhao et al. (2018) fused the features extracted from a pre-trained CNN and a fully-connected neural network to jointly perform eye state classification. However, these two deep models did not attempt to link the eye states with driver fatigue states. To address this problem, Chirra et al. (2019) predicted fatigue state by detecting 10 continuous frames of closed eye. More researchers (Hao et al., 2019; Liu et al., 2019c; Zhang et al., 2017) combined the PERCLOS criterion with CNN to predict

Table 5
Performance of eye state classification by different algorithms.

Algorithm	Accuracy
SVM	90.37%
Random forest	94.76%
MLP	95.18%
CNN, Kim et al. (2017)	Equal Error Rate = 0.2366
CNN + transfer learning, Zhao et al. (2018)	97.19%

the fatigue state. Xiao et al. (2019) proposed an end-to-end model by replacing the PERCLOS criterion with an LSTM network and achieved better accuracy. However, this method is very time-consuming.

As shown in Table 5, these deep learning based methods can achieve better performance compared to traditional algorithms such as SVM, Random Forest, and MLP.

Gaze zone feature. Besides the eye state, the eye gaze zone can also reflect the driver's fatigue state. An good gaze zone estimation model should simultaneously take different factors into account, such as eye features, relative face position, and head pose features. Zhang et al. (2015c) combined the eye features extracted by CNN with face angle calculated by a 3D shape model. The iTracker (Krafka et al., 2016) extracted eye and face features both from Convolutional Neural Network (CNN), and also combine the face position information. Though combining different factors can improve the accuracy, it increases the complexity and consumes more computing time. In Choi et al. (2016), researchers accomplished real-time gaze zone classification (with an average 24 fps) by using the eye region patches alone. George and Routray (2016) took the whole face image as the input, but the processing speed dropped to 20 fps. Moreover, in those real-time methods (Choi et al., 2016; George and Routray, 2016), the gaze zone estimation was regarded as a classification task instead of a regression task (Zhang et al., 2015c; Krafka et al., 2016). Therefore, the trade-off between time complexity and accuracy is an important issue for gaze zone estimation models.

Yawning feature. As another important behavior feature, the yawning feature is valued by many researchers for driver fatigue detection. Zhang et al. (2015b) proposed a deep learning-based driver yawning detection model consisting of a face detector, a nose detector, a nose tracker, and a yawning detector. According to the extracted features from a single image, a neural network is used to output a confidence score of yawning. Nevertheless, yawning is a continuous behavior that could last for dozens of frames. Detecting yawning from a single static image does not make full use of the contextual information. Therefore, more researchers used temporal deep models, such as the Long Short-Term Memory (LSTM) neural networks, to recognize driver's yawn.

Zhang et al. (2017) proposed to stack LSTM layers on the GoogLeNet to perform driver yawning detection. This model was modified by replacing LSTM with Bi-LSTM (Saurav et al., 2019). Xie et al. (2018) first trained an image-based yawning detection model, and then converted the model into a video-based model by combining LSTM. To better recognize short-term subtle facial action for yawning detection, Yang et al. (2021) combined Bi-LSTM with a 3D-CNN, and added a key frame selection step to remove the redundant frames and enhance the model efficiency for real-time performance.

Head pose feature. Deep learning-based head pose estimation emerged in 2014. Ahn et al. (2014) trained CNN to directly output a 3-D head pose vector, a Bayesian sequential estimation module was coupled to avoid jittery output. To further capture the contextual information, Borghi et al. (2017) stacked LSTM layers on CNN. The input images of the model were captured by a depth camera, while in Venturelli et al. (2017), two depth cameras at different angles were deployed to obtain multi-view visual information and achieve higher accuracy. Moreover, to solve the problem of lacking large amount of annotated training data, Liu et al. (2016) generated virtual samples with variations by a 3D shape model. Different from the above image-based head pose estimation models, Han et al. (2019a) used a neural network and MEMS magnetometer attaching to driver's neck to perform head pose estimation. Yang et al. (2020) applied Radio Frequency Identification (RFID) sensors to detect driver's head movement. They used an Variation Auto-Encoder (VAE)-based anomaly detection method to detect nodding behaviors. However, compared with the above non-invasive approaches, these two methods are less user-friendly.

Latent feature. The above methods were designed to detect fatigue state by focusing on specific fatigue features, such as eye state, gaze zone, head pose, etc. Some researchers adopted a different strategy by trying to extract latent fatigue-related feature and combine it with softmax classifiers (Dwivedi et al., 2014; Koesdwiady et al., 2017a; Park et al., 2016), or traditional classifiers such as SVM (Lopez et al., 2017) and Gradient Boosting Machine (Huynh et al., 2016). In addition, many researchers have discovered the potential of local methods, so parallel CNNs can be deployed to simultaneously take the global face and local patches as input to extract different levels of features (Reddy et al., 2017; Lyu et al., 2018; Huang et al., 2020; Liu et al., 2019b).

Most of the above-mentioned methods (Dwivedi et al., 2014; Koesdwiady et al., 2017a; Park et al., 2016; Reddy et al., 2017) used a single image as the input without considering the contextual information, while others extracted video-based features. Zhao et al. (2017) performed latent fatigue feature extraction by a video-based DBN model. In order to extract temporal features and spatial features simultaneously, Huynh et al. (2016) developed a 3D CNN model. Ma et al. (2019a) further proposed a model with three parallel 3D CNNs, which respectively take original image sequence, optical-flow images, and Optical Flow Motion History Images (OF-MHI) as inputs. However, the time window for 3D CNN model is fixed, the model might fail when fatigue behaviors have longer duration. In contrast, the LSTM model is capable of both short-term and long-term recognition. Models combining one, two, and three LSTM layer(s) with CNN have been reported to extract spatial–temporal features for fatigue detection (Shih and Hsu, 2016; Huang et al., 2020; Lyu et al., 2018).

4.2.3. Vehicle driving features

Vehicle-related applications, such as drive scene perception, path planning, and lane detection, have been powered by deep learning. Kim and Lee (2014) proposed to combine CNN and random sample consensus to perform robust lane detection. Lee et al. (2017) designed a multi-task structured Vanishing Point Guided Network (VPGNet). Experiments (Feng et al., 2018) show that VPGNet outperformed object detection model R-FCNN (Dai et al., 2016). Recently, Zou et al. (2019) combined LSTM with CNN to perform video-based lane detection. Compared with inferring lane position on a single image, video-based methods can achieve higher accuracy and handle difficult situations. However, these methods did not use extracted features to infer fatigue state. Vehicle-related applications, when handcrafted features were used, neural network-based methods (Zhao et al., 2019; Li et al., 2017; Chen et al., 2015) demonstrated sound capability on fatigue state prediction. Li et al. (2019b) showed that RNN with a fuzzy layer can also learn fatigue features in the time domain. These methods are all based on steering wheel angle or yaw angle features which can be easily collected through sensors. However, they are also easily disturbed by the variations of driver's habits and traffic conditions.

4.2.4. Information fusion

Predicting driver's fatigue state is influenced by various types of information. The relationship between information can be used to jointly predict driver fatigue and complement each other if some sensors fail. As shown in Table 6, information fusion based deep fatigue detection models can be divided into four types: driver physiological signal fusion, driver behavior feature fusion, vehicle driving feature fusion, and multi-modal feature fusion.

Table 6
Information fusion based deep learning methods.

Type	Ref.	Fusion method	Fused features
Driver physiological signal fusion	Zhang et al. (2016)	ANN	EEG, EOG
	Huo et al. (2016)	ANN	EEG, EOG
	Wu et al. (2021)	ANN	EEG, EOG
	Li et al. (2018)	ANN	EEG, EOG
	Du et al. (2017)	Deep auto-encoder	EEG, EOG
	Zhang and Etemad (2019)	Capsule attention module	EEG, EOG
	Jiao et al. (2020)	LSTM	EEG, EOG
	Zhang et al. (2013)	ANN	EEG, EOG, EMG
	Han et al. (2020)	ANN	EEG, ECG, Respiration, EDA
Driver behavior feature fusion	Ji et al. (2019)	Rule-based	Eye State, Mouth
	Guo and Markoni (2019)	Binary coding	Eye State, Mouth
	Sikander and Anwar (2021)	CNN	Eye State, Mouth, Forehead
	Yu et al. (2016)	CNN	Eye State, Mouth, Head
	Zhang et al. (2019)	ANN	Eye State, Mouth, Head
	Celona et al. (2018)	Attention module	Eye State, Mouth, Head
	Weng et al. (2016)	HMM	Eye State, Mouth, Head
Vehicle driving feature fusion	Tango and Botta (2013)	ANN/RNN	Speed, Time to collision, Time to lane crossing, Steering angle, Lateral position, Position of the accelerator pedal, Position of the brake pedal.
	Shahverdy et al. (2020)	CNN	Acceleration, Gravity, Throttle, Speed, RPM
Multi-modal feature fusion	Utomo et al. (2019)	LSTM	Eye State, Heart Rate
	Li et al. (2019a)	Active Learning	Eye State, Mouth, SWA
	Jain et al. (2016)	ANN	Vision Feature, Road Facing Camera Images
	Streiffer et al. (2017)	ANN	Vision Feature, IMU data
	Karuppusamy and Kang (2020)	ANN	EEG, Accelerometer, Gyroscope Sensor, Vision Feature
	Lim and Yang (2016)	CNN	Velocity, Acceleration, Steering Wheel Angle, Gas Pedal Angle, Blinking Rate, PERCLOS, Facial Direction, Audio Amplitudes, Heart Rate, Respiration Rate, Galvanic Skin Response, Body Temperature

Driver physical feature fusion is normally done on EEG and EOG signals. Several deep models, including fully-connected neural network, deep auto-encoder, and capsule attention module, have been used to fuse them (Zhang et al., 2016; Huo et al., 2016; Wu et al., 2021; Li et al., 2018; Du et al., 2017; Zhang and Etemad, 2019). Jiao et al. (2020) proposed to segment the EOG signal samples based on the EEG signal, and then fed the segmented EOG samples into the LSTM network to predict fatigue. They also used GAN-based data augmentation to alleviate the insufficient data problem. Moreover, by using EEG, EMG, and EOG signals, Zhang et al. (2013) proposed a real-time fatigue detection method based on various entropy and complexity measures. ECG, EDA, and respiration signals have also been combined to perform fatigue detection (Han et al., 2020).

In fusion of driver behavior features, highly distinguishable eye and mouth features are commonly used. For example, SP et al. (2021) combined face key points, eye and mouth opening states to inference fatigue state jointly. Ji et al. (2019) proposed eye state and mouth state recognition networks and fused multi-indicators for fatigue judgment. Guo and Markoni (2019) used CNNs to analyze the eyes states and mouth state, then used LSTM to extract temporal features from the visual information. Sikander and Anwar (2021) reconstructed 3D shapes of the eye, mouth, and forehead and calculated their quiver map that represent subtle facial action. Many approaches also used head pose features in the prediction (Yu et al., 2016; Zhang et al., 2019; Celona et al., 2018; Weng et al., 2016). The relevance of multiple features, drivers status and their temporal dependencies can be further modeled using attention module (Celona et al., 2018) and Hidden Markov Model (HMM) (Weng et al., 2016).

Tango and Botta (2013) fused various types of vehicle information including speed, time to collision, time to lane crossing, steering angle, lateral position, position of the accelerator pedal, and position of the brake pedal. A fully-connected neural network and an RNN were utilized to analyze the input features. Shahverdy et al. (2020) converted vehicle signals (including acceleration, gravity, throttle, speed, and revolutions per minute) to a 2-D matrix by calculating the recurrence plot, then used CNN for feature extraction and classification.

The above mentioned fusion methods are all based on the same type of information which may be highly correlated to each other or cannot provide complementary information. For example, both deep visual features of eye state and yawn detection may fail when the vehicle is in a dark environment, causing corresponding fusion failure. In contrast, multi-modal fusion methods could benefit from highly complementary information. Omerustaoglu et al. (2020) proved this point by comprehensive comparative experiment. They found that compared to using vision-only information, further utilization of sensor data (e.g., speed, fuel level, acceleration, etc.) can improve the model performance by up to 9%. Utomo et al. (2019) used PERCLOS and heart rate features as the input to the LSTM network. Li et al. (2019a) calculated parameters including eye open-angle, mouth open-angle, change of the SWA, and the speed of SWA change. These parameters were concatenated into a vector and inputted into a semi-supervised active learning model. By adding a fully-connected layer in the LSTM model, Jain et al. (2016) fused vision features and vehicle driving features. In the DarNet proposed by Streiffer et al. (2017), vision features extracted by CNN and IMU data collected from the driver's mobile device were fused. Moreover, features in all the three modalities, including physiological signal feature, behavior feature, and vehicle driving feature, can be combined by parallel deep networks (Karuppusamy and Kang, 2020) and CNN (Lim and Yang, 2016). Zhang et al. (2021) fused vehicle sensors signal, an acoustic signal and visual signal, and then developed

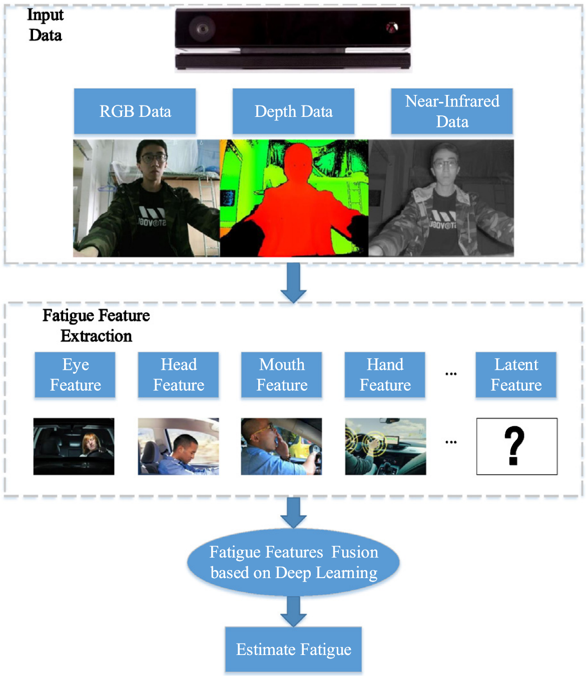

Fig. 9. Illustration of driver fatigue detection based on the integration of RGB-D camera and deep learning.

a driver distraction detection method based on a reconstruction-based anomaly detection scheme. Driver distraction and driver fatigue differ in the fact that distracted driver could make additional sounds, but fatigue driver is usually quiet. Therefore the acoustic signal adopted in Zhang et al. (2021) may not be useful for driver fatigue detection.

5. Integration of RGB-D camera and deep learning

Based on the above analysis, it is natural to consider integration of RGB-D camera and deep learning technologies. The integration framework is shown in Fig. 9. First, by making full use of RGB images, depth images, and near-infrared images, those observable fatigue features such as hand behavior, head posture, eye state, and mouth state can be effectively extracted. In addition, the latent driver fatigue features which has powerful representation capabilities can be extracted by a deep learning model. Finally, the extracted fatigue features are fused to estimate driver's fatigue state.

Integration of RGB-D camera and deep learning will bring two benefits. First, the depth image and near-infrared image from RGB-D camera can be used as supplements to the standard RGB cameras. For example, depth image and near-infrared image are effective in both day and night environments, but RGB image has better resolution. The combination of them can improve the accuracy and robustness of the fatigue detection system. Second, the representation capability of deep latent features is much better than traditional hand-craft features, leading to higher fatigue detection accuracy. Moreover, unlike traditional methods such as Cyganek and Gruszczyński (2014), the deep learning model can automatically learn useful information across different channels. Therefore it can fuse the information from RGB, depth, and near-infrared images more effectively. In the rest of this section, we present a set of experiments to validate the effectiveness of this framework from different aspects.

In this section, we will conduct a series of experiments to validate this framework. In Section 5.1, we first compare deep features with traditional visual descriptors and show its superiority, then compare different deep features and found that CNN achieves the best performance. Then, we demonstrate the data hungriness of deep features, and show that a GAN-based data augmentation strategy can alleviate this issue in Section 5.2. Finally in Section 5.3, we collect a driver fatigue detection dataset and conduct experiments with a variety of learning algorithm to demonstrate the effectiveness of integrating RGB-D camera and deep learning.

Table 7
Accuracies of CNN and hand-crafted features for eye state classification.

Eye feature	Accuracy
Local Binary Pattern (LBP) (Ojala et al., 2002)	92.0%
Weber Local Descriptor (WLD) (Chen et al., 2010)	94.5%
Weber Local Binary Pattern (WLBP) (Liu et al., 2013)	96.6%
Convolutional Neural Network (CNN)	**98.2%**

5.1. Effectiveness of deep features

We first demonstrate the advantage of CNN over traditional methods through the eye state classification task. Three typical traditional handcrafted features including Local Binary Pattern (LBP) (Ojala et al., 2002), Weber Local Descriptor (WLD) (Chen et al., 2010), and Weber Local Binary Pattern (WLBP) (Liu et al., 2013) are compared with CNN. In the experiment, 7686 images from RPI ISL Eye Database (Wang and Ji, 2005) containing 3456 open eye images and 4140 close eye images are adopted. We use 1576 opened eye images and 1840 closed eye images for training and use the remaining images for testing. All the eye images are normalized to the size of 28×28. For the handcrafted features, SVM (Chang and Lin, 2011) is used as classifier. For the CNN features, the structure of CNN model is shown in Fig. 12. The testing accuracies are shown in Table 7. Not surprising, the performance of CNN is better than LBP, WLD, and WLBP features.

Besides these observable fatigue features, we also extract and compare the latent fatigue features by three deep learning technologies CNN (Krizhevsky et al., 2012), Deep Belief Nets (DBN) (Hinton et al., 2006), and Stacked Denoising AutoEncoders (SDAE) (Vincent et al., 2010). We first construct a dataset containing 4441 face images, in which 2000 images are labeled as fatigue state by professional data annotator and the others are regarded as awake state. In the constructed dataset, 543 face images are collected by a standard RGB-D camera, and 3898 face images are extracted from public face datasets including AR (Martínez, 1998), YaleA (Belhumeur et al., 1997), YaleB (Georghiades et al., 2001), and JAFFE (Lyons et al., 1998). The train-test split ratio is 8:2. Fig. 10 shows the training losses and testing accuracies of CNN, DBN, and SDAE. Compared with DBN and SDAE, CNN achieves the best fatigue detection result. It demonstrates the effectiveness of CNN and its advantage over DBN and SDAE. Therefore, in the subsequent experiments, we will utilize CNN to extract the latent fatigue feature.

5.2. Effectiveness of data augmentation

Although CNN achieves the best results, we find that its performance heavily relies on the number of training samples. To demonstrate this point, we train a series of CNN with different sizes of the training set, then compare their accuracies. The training samples are randomly selected, and the remaining samples are utilized for testing. As shown in Fig. 11, when the number of training samples is limited to 500, the accuracy dropped to 80%. When there are only 10 training samples, the accuracy is slightly higher than the random guess (50%). It shows that the performance of deep learning methods are significantly influenced by the size of training set. But in some cases, it is difficult yet expensive to construct large-scale annotated dataset for driver fatigue detection. Therefore, it is necessary to consider data augmentation to further improve the performance.

Based on the comparison between CNN and traditional features in the last section, we can attribute the advantage of CNN over traditional

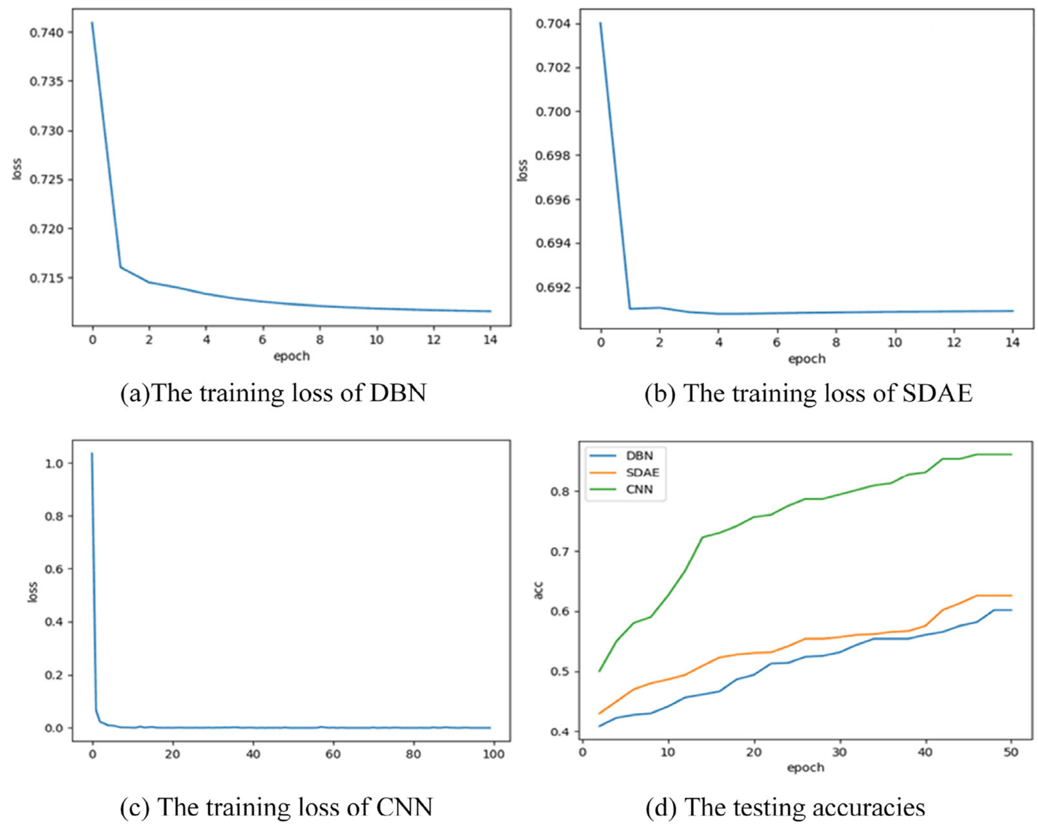

Fig. 10. The training loss and testing accuracies of DBN, SDAE, and CNN for latent fatigue detection.

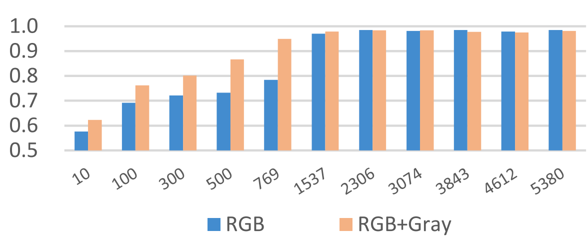

Fig. 11. Eye state detection accuracies of three-channel (RGB) and four-channel (RGB-D) CNN with different number of training samples.

features to the ability of learning pixel-wise patterns. Except this, CNN may also have the potential to learn complementary channel-wise information. This property could benefit the model when taking multiple-source inputs, such as depth, near-infrared and standard RGB channels. To verify this point, we test a four-channel scheme for data augmentation, where an additional gray-scale channel is integrated with the RGB channel. As shown in Fig. 11, the four-channel scheme is superior to the three-channel, especially when the number of training sample is limited. The improvement from adding channel demonstrates that CNN successfully learns to extract complementary channel-wise information. Therefore, we believe that the multi-channel scheme of CNN is an effective way to integrate the multi-source information such as RGB image and depth image.

One of the most widely used deep learning technologies for data augmentation is GAN (Goodfellow et al., 2014). Compared to random scale and crop, GAN-based models can automatically learn the distribution of real data and generate realistic sample. However, there are few works using GAN for driver fatigue detection. To validate its feasibility, we use the Deep Convolutional Generative Adversarial Networks (DCGAN) (Radford et al., 2016) to generate more virtual eye image samples for training the CNN, which is illustrated in Fig. 12. The generated eye samples are shown in Fig. 13. To validate its effectiveness, the training samples from the original training set are first used to train the DCGAN model to generate new samples. Then, we use the trained DCGAN to generate 10, 100, 300, 500, 700, 1000, 2000, 3000 samples and mix them with the original training samples to obtain a series of augmented training sets. These augmented training sets are respectively utilized for training CNNs. The test results with three-channel and four-channel CNN are shown in Figs. 14 and 15, respectively, where n represents the size of the original dataset. It can be seen that the augmented samples from DCGAN can effectively improve the classification accuracy.

Table 8
Classification accuracies based on RGB-D camera for different tasks.

Task		Accuracy
Eye	Open/closed	89%
Mouth	Open/closed	62%
Head	Normal/drop	67%

5.3. Effectiveness of RGB-D camera and information fusion

To validate the effectiveness of the RGB-D camera, we compare the reliability of eye, mouth and head fatigue features extracted from Kinect. For eye and mouth behaviors, we respectively test their recognition accuracy of open and closed states. For head, we test the recognition accuracy of normal and head drop states. The experiment for each task is conducted ten times independently, and the average results are reported in Table 8. It shows that RGB-D camera is an effective way to extract different fatigue features. However, the mouth and head features are less reliable than eye features. Therefore, the fusion of different fatigue features is necessary to get more robust results.

To demonstrate the effectiveness of information fusion, we conduct comparative experiment to validate different types of fatigue features.

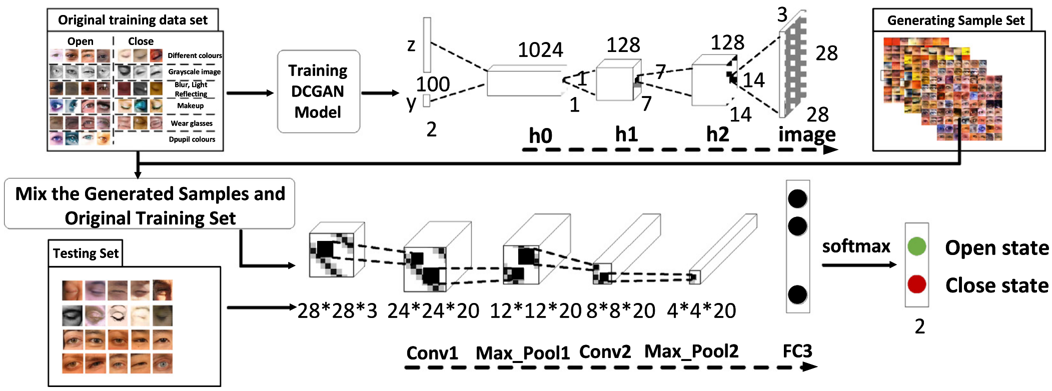

Fig. 12. Illustration of DCGAN+CNN for eye state detection. The training set contains various kinds of eye images such as different skin colors, gray images, blur, light-reflecting, makeup, different pupil colors, and so on.

Fig. 13. Generated eye samples by DCGAN.

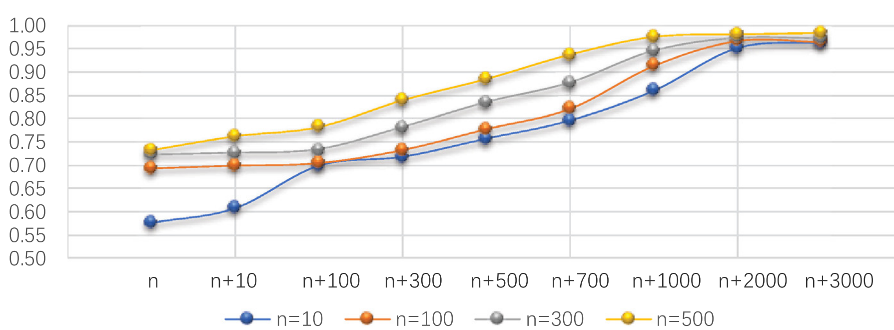

Fig. 14. Testing accuracies under different number of original training samples and augmented samples by three-channel DCGAN+CNN.

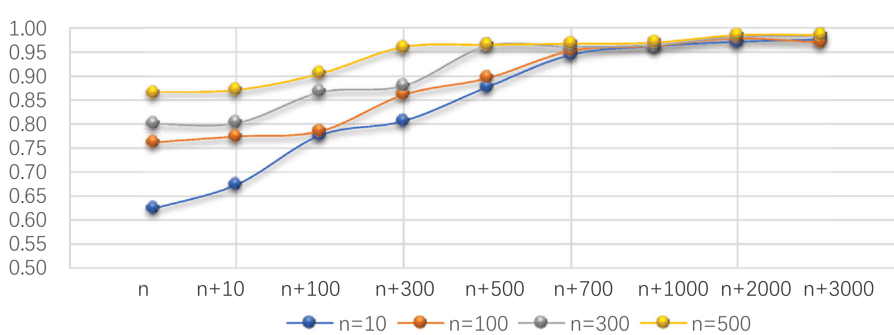

Fig. 15. Testing accuracies under different numbers of original training samples and augmented samples by four-channel DCGAN+CNN.

We collect a driver fatigue video dataset[1] of 17 drivers containing fatigue state or non-fatigue state, where the videos of non-fatigue state are collected when the drivers have sufficient sleeping, while fatigue state videos are collected under sleep deprivation condition. Some samples in this dataset are shown in Fig. 16. Each video clip lasts about 100 s. We use a 300-frame sliding-window and a stride of 5 frames to collect training and testing samples. Each sample consists of 300

[1] The dataset is available at https://github.com/ChenDelong1999/DriverFatigueDetection.

Fig. 16. Samples of the collected dataset. Upper row: fatigue state. Bottom row: non-fatigue state.

Table 9
Representation formats and the fusion strategies.

Feature		Value	Fusion1	Fusion2
Eye	Left Right	Unknow/Opened/Maybe/Closed	✓	✓
Mouth	Closure Moving	Unknow/No/Maybe/Yes	✓	✓
Head		Yaw, Pitch, Roll	✓	✓
Emotion		Unknow/Not happy/Maybe/happy	✓	✓
Latent (ResNet He et al., 2016)		Fatigue/Non-fatigue		✓

Table 10
Results of fatigue detection using different features, classifiers, and fusion strategies.

	Eye	Mouth	Head	Fusion1	Fusion2
KNN	71.90%	63.76%	63.61%	67.27%	69.67%
Decision Tree	81.52%	56.75%	59.12%	80.82%	83.01%
AdaBoost	82.02%	69.37%	**72.34%**	86.80%	89.06%
Logistic regression	84.04%	62.36%	64.87%	85.67%	87.60%
Random forest	89.63%	61.77%	62.04%	88.88%	89.76%
Gradient boosting	90.83%	**73.73%**	64.47%	92.14%	92.48%
SVM	90.33%	70.12%	64.87%	87.75%	90.70%
LSTM	**92.00%**	69.00%	61.00%	**93.00%**	**94.00%**
Average	85.29%	65.86%	64.04%	85.29%	87.03%

video frames and a label indicating the corresponding fatigue state. The dataset are divided into training set and testing set whose sizes are respectively 13,402 and 7332.

In our experiments, eye state, mouth state, head pose and face emotion state estimated by the RGB-D camera are used as fatigue feature. The eye state, mouth state and face emotion predictions are converted to two-way one-hot vectors, while head pose is represented as 3-dimensional head orientation vector. Moreover, a standard ResNet-18 (He et al., 2016) is trained from scratch to extract latent fatigue feature. The ResNet-18 backbone takes normalized 224 × 224 RGB image as input. A linear classifier is attached to the 1024-dimensional feature to convert it to two-way class possibility scores. The network is trained with cross-entropy loss function by stochastic gradient descent (SGD) with a learning rate of 10^{-3} for 100 epochs. To verify the advantages of information fusion, two types of fusing strategies were tried. The first one is to fuse all types of fatigue features, and the other does not consider latent fatigue features. These different feature representation and fusion strategies are summarized in Table 9.

As shown in Table 10, we respectively use KNN, decision tree, logistic regression, random forest, gradient boosting, SVM, and LSTM to detect the fatigue state based on different fatigue features. The experiment results clearly show that the eye features are more effective than the mouth and the head pose features. Moreover, the fused features outperform any type of single feature. LSTM achieves the best result by integrating all kind of fatigue features. It clearly demonstrates the advantage of combining RGB-D camera and deep learning.

5.4. Discussion and future research directions

The goal of the above experiments is to demonstrate the effectiveness of integrating RGB-D camera and deep learning from different perspectives. CNN-based methods have outperformed several traditional methods, and its combination with RGB-D camera can further enhance the performance. Therefore, deep learning and multi-sensor information fusion have become an increasingly important research trend on driver fatigue detection. We further demonstrated that large-scale training data is an important prerequisite of developing a deep learning-based driver fatigue detection system. When the training data is not sufficient to enable the CNN to learn generalizable representations, the fatigue detection accuracy will drop dramatically due to the overfitting problem.

However, it is non-trivial to construct a driver fatigue detection dataset that is not only large-scale, but also multi-modal and high-quality. Due to the lack of a well-recognized public benchmark dataset, current fatigue detection research is mostly undertaken on non-public datasets, which greatly limits the fair evaluation of different approaches. The difficulty of obtaining training data lies in several aspects. First, the cost of constructing a dataset is high, no matter based on driving simulator or real driving environment. A driving simulator cannot fully simulate driving conditions, including acceleration feedback, the influence of passengers, and other in-car environments, etc. On the other hand, collecting a driver fatigue dataset in a real environment brings additional safety risks and ethical issues (Němcová et al., 2020). Finally, the annotation of fatigue status is often based on subjective assessment and lacks objectivity.

A promising solution to the above problem is transfer learning. Standard transfer learning includes two steps: (1) pretraining on a large-scale dataset for learning general knowledge and (2) fine-tuning on a target dataset for learning domain-specific knowledge. A recent from Omerustaoglu et al. (2020) found that loading the CNNs pretrained on ImageNet and fine-tuning them on the driver fatigue detection task is beneficial. In addition to deep models for image classification, transferring deep models pre-trained on EEG classification (Bigdely-Shamlo et al., 2016), face emotion recognition (Ben et al., 2021), or vehicle trajectory prediction (Xing et al., 2020) tasks have great potential to enable the deep model to learn semantic feature representations with a limited training set. However, there are several points that need additional attention when employing transfer learning. First, public available pretrained CNNs are usually deep and large, while the driver fatigue detection dataset sizes are usually quite small. Fine-tuning large models on a small dataset has the potential of the "catastrophic forgetting" of learned knowledge from pretraining due to severe over-fitting. Besides, training samples of driver fatigue detection dataset are usually very different compared to images in general image classification dataset. The domain gap between the pretraining dataset and the downstream dataset needs to be taken into consideration. Introducing domain adaptation (Wang and Deng, 2018) techniques can possibly address this issue.

In addition, recent advances and success of multi-modal self-supervised learning (Arandjelovic and Zisserman, 2017; Korbar et al., 2018; Patrick et al., 2020) implies the potential of learning fatigue feature relationships from large-scale unlabeled vehicle record data and in-vehicle videos. From another perspective, transfer learning and self-supervised learning can also be regarded as introducing prior knowledge from supernumerary data, and consequently alleviate the limitation of the training set size on the generalization ability. Therefore, when encountering the lack of large-scale data, introducing transfer learning and multi-modal self-supervised learning are promising future research directions for driver fatigue detection.

Moreover, privacy issues need to be taken into consideration for these computer-vision-based approaches. Most existing approaches primarily focus on improving driver fatigue detection performance, but very few of them have considered the privacy issue. For example, model compression via knowledge distillation can be applied to make the deep models more computationally efficient. If fatigue detection model are scale down to proper size while maintaining high performance, it can be deployed locally on in-car devices, without any communication with servers through the Internet.

6. Conclusion

Effective fatigue feature extraction is the key point in fatigue detection. However, differences in fatigue characteristics produced by different subjects and different driving environments limit the generality of single type of fatigue feature. In this review, we studied and analyzed the advantages and disadvantages of various fatigue detection methods, which show that RGB-D camera and deep learning are two promising directions for future research. Multi-source data from a RGB-D camera contains valuable information to ensure the effectiveness of fatigue detection. This can be boosted by deep learning technologies which capture various observable and latent fatigue characteristics. Therefore, we expect using RGB-D camera and deep learning technology simultaneously will be the potential new developing trend. We undertook some baseline experiments in this direction, which verify the advantages of these two technologies. The experiments also demonstrates that information fusion technology can significantly decrease the dependence on the specific single fatigue features. In general, the above three techniques are worthwhile of further investigation for fatigue detection, which will effectively promote robustness and reliability on the task.

CRediT authorship contribution statement

Fan Liu: Conceptualization, Methodology, Writing, Funding acquisition. **Delong Chen:** Writing, Editing, Validation. **Jun Zhou:** Review and editing. **Feng Xu:** Review and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially funded by Natural Science Foundation of Jiangsu Province, China under grant No. BK20191298, Research Fund from Science and Technology on Underwater Vehicle Technology Laboratory (2021JCJQ-SYSJJ-LB06905), Water Science and Technology Project of Jiangsu Province under grant No. 20210722021063, and the Qinglan Project of Jiangsu Province.

References

Abbas, Q., Alsheddy, A., 2021. Driver fatigue detection systems using multi-sensors, smartphone, and cloud-based computing platforms: a comparative analysis. Sensors 21 (1), 56.

2020. ADAS-P9 Automobile Anti-collision Warning System. [EB/OL], http://www.cheyuansu.com.cn/adas_products_8.html Accessed December 1, 2020.

Ahn, S., Nguyen, T., Jang, H., Kim, J.G., Jun, S.C., 2016. Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. Front. Human Neurosci. 10, 219.

Ahn, B., Park, J., Kweon, I.S., 2014. Real-time head orientation from a monocular camera using deep neural network. In: Asian Conference on Computer Vision. Springer, pp. 82–96.

Al-Sultan, S., Al-Bayatti, A.H., Zedan, H., 2013. Context-aware driver behavior detection system in intelligent transportation systems. IEEE Trans. Veh. Technol. 62 (9), 4264–4275.

Alsibai, M.H., Manap, S.A., 2016. A study on driver fatigue notification systems. ARPN J. Eng. Appl. Sci. 11 (18), 10987–10992.

Arandjelovic, R., Zisserman, A., 2017. Look, listen and learn. In: 2017 IEEE International Conference on Computer Vision. pp. 609–617. http://dx.doi.org/10.1109/ICCV.2017.73.

2020. Asleep at the Wheel: A Nation of Drowsy Drivers. [EB/OL], https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/12723-drowsy_driving_asleep_at_the_wheel_031917_v4b_tag.pdf. Accessed December 1, 2020.

Bani, I., Akrout, B., Mahdi, W., 2019. Real-time driver fatigue monitoring with a dynamic Bayesian network model. In: Digital Health Approach for Predictive, Preventive, Personalised and Participatory Medicine. Springer, pp. 69–77.

Belhumeur, P., Hespanha, J., Kriegman, D., 1997. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. 19, 711–720.

Ben, X., Ren, Y., Zhang, J., Wang, S.-J., Kpalma, K., Meng, W., Liu, Y.-J., 2021. Video-based facial micro-expression analysis: A survey of datasets, features and algorithms. IEEE Trans. Pattern Anal. Mach. Intell. 1.

Bergasa, L., Nuevo, J., Sotelo, M., Barea, R., Lopez, M., 2006. Real-time system for monitoring driver vigilance. IEEE Trans. Intell. Transp. Syst. 7 (1), 63–77.

Bigdely-Shamlo, N., Makeig, S., Robbins, K.A., 2016. Preparing laboratory and real-world EEG data for large-scale analysis: A containerized approach. Front. Neuroinform. 10, 7.

Borghi, G., Gasparini, R., Vezzani, R., Cucchiara, R., 2017. Embedded recurrent network for head pose estimation in car. In: IEEE Intelligent Vehicles Symposium. IEEE pp. 1503–1508.

Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., Candefjord, S., 2019. Deriving heart rate variability indices from cardiac monitoring—An indicator of driver sleepiness. Traffic Inj. Prev. 20 (3), 249–254.

Cai, Q., Gao, Z.-K., Yang, Y.-X., Dang, W.-D., Grebogi, C., 2019. Multiplex limited penetrable horizontal visibility graph from EEG signals for driver fatigue detection. Int. J. Neural Syst. 29 (05), 1850057.

Cajochen, C., Foy, R., Dijk, D.-J., et al., 1999. Frontal predominance of a relative increase in sleep delta and theta EEG activity after sleep loss in humans. Sleep Res. Online 2 (3), 65–69.

Calcagnini, G., Biancalana, G., Giubilei, F., Strano, S., Cerutti, S., 1994. Spectral analysis of heart rate variability signal during sleep stages. In: Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 2. IEEE, pp. 1252–1253.

Cao, Y., Lu, B.-L., 2013. Real-time head detection with kinect for driving fatigue detection. In: International Conference on Neural Information Processing. Springer, pp. 600–607.

Cecotti, H., Graser, A., 2010. Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE Trans. Pattern Anal. Mach. Intell. 33 (3), 433–445.

Celona, L., Mammana, L., Bianco, S., Schettini, R., 2018. A multi-task CNN framework for driver face monitoring. In: IEEE 8th International Conference on Consumer Electronics-Berlin. IEEE, pp. 1–4.

Chacon-Murguia, M.I., Prieto-Resendiz, C., 2015. Detecting Driver Drowsiness: A survey of system designs and technology. IEEE Consum. Electron. Mag. 4 (4), 107–119.

Chang, T.-H., Hsu, C.-S., Wang, C., Yang, L.-K., 2008. Onboard measurement and warning module for irregular vehicle behavior. IEEE Trans. Intell. Transp. Syst. 9 (3), 501–513.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) 2 (3), 1–27.

Chen, J., Shan, S., He, C., Zhao, G., Pietikäinen, M., Chen, X., Gao, W., 2010. WLD: A robust local image descriptor. IEEE Trans. Pattern Anal. Mach. Intell. 32 (9), 1705–1720.

Chen, H., Xue, Y., Xu, C., Wang, X., Pei, L., Han, Y., 2018. An algorithm to determine driver's fatigue driving according to characteristic parameters based on the theory of rough set importance. In: The 3rd International Conference on Automation, Mechanical Control and Computational Engineering. Atlantis Press, pp. 334–337.

Chen, Z., Yang, P., Peng, L., Mo, Z., Cai, G., 2015. Fatigue driving monitoring based on BP neural network. Comput. Sci. 42 (Z6), 67–69.

Chhabra, R., Krishna, C.R., Verma, S., 2019. Smartphone based context-aware driver behavior classification using dynamic bayesian network. J. Intell. Fuzzy Systems 36 (5), 4399–4412.

Chirra, V.R.R., ReddyUyyala, S., Kolli, V.K.K., 2019. Deep CNN: A machine learning approach for driver drowsiness detection based on eye state. Rev. Intell. Artif. 33 (6), 461–466.

Choi, I.-H., Hong, S.K., Kim, Y.-G., 2016. Real-time categorization of driver's gaze zone using the deep learning techniques. In: International Conference on Big Data and Smart Computing. IEEE, pp. 143–148.

Choi, I.-H., Kim, Y.-G., 2014. Head pose and gaze direction tracking for detecting a drowsy driver. In: 2014 International Conference on Big Data and Smart Computing. IEEE, pp. 241–244.

Craye, C., Karray, F., 2015. Driver distraction detection and recognition using RGB-D sensor. arXiv preprint arXiv:1502.00250.

Cyganek, B., Gruszczyński, S., 2014. Hybrid computer vision system for drivers' eye recognition and fatigue monitoring. Neurocomputing 126, 78–94, Recent trends in Intelligent Data Analysis Online Data Processing.

Dai, J., Li, Y., He, K., Sun, J., 2016. R-FCN: Object detection via region-based fully convolutional networks. arXiv preprint arXiv:1605.06409.
Dingus, T.A., Jahns, S.K., Horowitz, A.D., Knipling, R., 1998. Human factors design issues for crash avoidance systems. Human Factors Intell. Transp. Syst. 55–93.
Dong, W., Qu, P., Han, J., 2008. Driver fatigue detection based on fuzzy fusion. In: Chinese Control and Decision Conference. IEEE, pp. 2640–2643.
Du, Y., Hu, Q., Chen, D., Ma, P., 2011. Kernelized fuzzy rough sets based yawn detection for driver fatigue monitoring. Fund. Inform. 111 (1), 65–79.
Du, G., Li, T., Li, C., Liu, P.X., Li, D., 2021. Vision-based fatigue driving recognition method integrating heart rate and facial features. IEEE Trans. Intell. Transp. Syst. 22 (5), 3089–3100.
Du, L.-H., Liu, W., Zheng, W.-L., Lu, B.-L., 2017. Detecting driving fatigue with multimodal deep learning. In: The 8th International IEEE/EMBS Conference on Neural Engineering. IEEE, pp. 74–77.
Du, G.L., Wang, Z., Li, C., Liu, P.X., 2020. A TSK-type convolutional recurrent fuzzy network for predicting driving fatigue. IEEE Trans. Fuzzy Syst. 1.
Dwivedi, K., Biswaranjan, K., Sethi, A., 2014. Drowsy driver detection using representation learning. In: IEEE International Advance Computing Conference. IEEE, pp. 995–999.
Eskandarian, A., Mortazavi, A., 2007. Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection. In: IEEE Intelligent Vehicles Symposium. IEEE, pp. 553–559.
2020. FaceLab 5. [EB/OL], https://www.eyecomtec.com/3132-facelab Accessed December 1, 2020.
Fan, X., Yin, B.-C., Sun, Y.-F., 2007. Yawning detection for monitoring driver fatigue. In: International Conference on Machine Learning and Cybernetics, Vol. 2. IEEE, pp. 664–668.
Feng, J., Wu, X., Zhang, Y., 2018. Lane detection base on deep learning. In: The 11th International Symposium on Computational Intelligence and Design, Vol. 1. IEEE, pp. 315–318.
Friedrichs, F., Yang, B., 2010. Camera-based drowsiness reference for driver state classification under real driving conditions. In: IEEE Intelligent Vehicles Symposium. IEEE, pp. 101–106.
Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., Zuo, S., 2019. EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. IEEE Trans. Neural Netw. Learn. Syst. 30 (9), 2755–2763.
García, F., de la Escalera, A., Armingol, J.M., et al., 2014. Driver monitoring based on low-cost 3-D sensors. IEEE Trans. Intell. Transp. Syst. 15 (4), 1855–1860.
George, A., Routray, A., 2016. Real-time eye gaze direction classification using convolutional neural network. In: International Conference on Signal Processing and Communications. IEEE, pp. 1–5.
Georghiades, A., Belhumeur, P., Kriegman, D., 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Mach. Intell. 23 (6), 643–660.
Golz, M., Sommer, D., Trutschel, U., Sirois, B., Edwards, D., 2010. Evaluation of fatigue monitoring technologies. Somnologie-Schlafforschung Schlafmed. 14 (3), 187–199.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A.C., Bengio, Y., 2014. Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada. pp. 2672–2680.
Gu, H., Ji, Q., Zhu, Z., 2002. Active facial tracking for fatigue detection. In: Th 6th IEEE Workshop on Applications of Computer Vision. IEEE, pp. 137–142.
Guo, J.-M., Markoni, H., 2019. Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. Multimedia Tools Appl. 78 (20), 29059–29087.
Hajinoroozi, M., Jung, T.-P., Lin, C.-T., Huang, Y., 2015. Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data. In: IEEE China Summit and International Conference on Signal and Information Processing. IEEE, pp. 812–815.
Hajinoroozi, M., Zhang, J.M., Huang, Y., 2017. Driver's fatigue prediction by deep covariance learning from EEG. In: IEEE International Conference on Systems, Man, and Cybernetics. IEEE, pp. 240–245.
Han, H., Jang, H., Yoon, S.W., 2019a. Driver head posture monitoring using mems magnetometer and neural network for long-distance driving fatigue analysis. In: IEEE SENSORS. IEEE, pp. 1–4.
Han, S.-Y., Kwak, N.-S., Oh, T., Lee, S.-W., 2020. Classification of pilots' mental states using a multimodal deep learning network. Biocybern. Biomed. Eng. 40 (1), 324–336.
Han, C., Sun, X., Yang, Y., Che, Y., Qin, Y., 2019b. Brain complex network characteristic analysis of fatigue during simulated driving based on electroencephalogram signals. Entropy 21 (4), 353.
Hao, Z., Wan, G., Tian, Y., Tang, Y., Dai, T., Liu, M., Wei, R., 2019. Research on driver fatigue detection method based on parallel convolution neural network. In: IEEE International Conference on Power, Intelligent Computing and Systems. IEEE pp. 164–168.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.
Healey, J., Picard, R., 2000. SmartCar: detecting driver stress. In: Proceedings 15th International Conference on Pattern Recognition, Vol. 4. pp. 218–221. http://dx.doi.org/10.1109/ICPR.2000.902898.
Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput. 18 (7), 1527–1554.
Hostens, I., Ramon, H., 2005. Assessment of muscle fatigue in low level monotonous task performance during car driving. J. Electromyogr. Kinesiol. 15 (3), 266–274.
Hu, J., Xu, L., He, X., Meng, W., 2017. Abnormal driving detection based on normalized driving behavior. IEEE Trans. Veh. Technol. 66 (8), 6645–6652.
Huang, R., Wang, Y., Li, Z., Lei, Z., Xu, Y., 2020. RF-DCM: Multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection. IEEE Trans. Intell. Transp. Syst..
Huo, X.-Q., Zheng, W.-L., Lu, B.-L., 2016. Driving fatigue detection with fusion of EEG and forehead EOG. In: International Joint Conference on Neural Networks. IEEE, pp. 897–904.
Huynh, X.-P., Park, S.-M., Kim, Y.-G., 2016. Detection of driver drowsiness using 3D deep neural network and semi-supervised gradient boosting machine. In: Asian Conference on Computer Vision. Springer, pp. 134–145.
Illingworth, J., Kittler, J., 1988. A survey of the Hough transform. Comput. Vis. Graph. Image Process. 44 (1), 87–116.
Isaza, C., Anaya, K., Fuentes-Silva, C., de Paz, J.P.Z., Rizzo, A., Garcia-Moreno, A.-I., 2019. Dynamic set point model for driver alert state using digital image processing. Multimedia Tools Appl. 78 (14), 19543–19563.
Jain, A., Singh, A., Koppula, H.S., Soh, S., Saxena, A., 2016. Recurrent neural networks for driver activity anticipation via sensory-fusion architecture. In: IEEE International Conference on Robotics and Automation. IEEE, pp. 3118–3125.
Jeong, I.C., Lee, D.H., Park, S.W., Ko, J.I., Yoon, H.R., 2007. Automobile driver's stress index provision system that utilizes electrocardiogram. In: IEEE Intelligent Vehicles Symposium. IEEE, pp. 652–656.
Ji, Y., Wang, S., Zhao, Y., Wei, J., Lu, Y., 2019. Fatigue state detection based on multi-index fusion and state recognition network. IEEE Access 7, 64136–64147.
Jiao, Y., Deng, Y., Luo, Y., Lu, B.-L., 2020. Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks. Neurocomputing 408, 100–111.
Kaplan, S., Guvensan, M.A., Yavuz, A.G., Karalurt, Y., 2015. Driver behavior analysis for safe driving: A survey. IEEE Trans. Intell. Transp. Syst. 16 (6), 3017–3032.
Karuppusamy, N.S., Kang, B.-Y., 2020. Multimodal system to detect driver fatigue using EEG, gyroscope, and image processing. IEEE Access 8, 129645–129667.
Kim, K.W., Hong, H.G., Nam, G.P., Park, K.R., 2017. A study of deep CNN-based classification of open and closed eyes using a visible light camera sensor. Sensors 17 (7), 1534.
Kim, J., Lee, M., 2014. Robust lane detection based on convolutional neural network and random sample consensus. In: International Conference on Neural Information Processing. Springer, pp. 454–461.
Knapik, M., Cyganek, B., 2019. Driver's fatigue recognition based on yawn detection in thermal images. Neurocomputing 338, 274–292.
Ko, W., Oh, K., Jeon, E., Suk, H.-I., 2020. VIGNet: A deep convolutional neural network for EEG-based driver vigilance estimation. In: The 8th International Winter Conference on Brain-Computer Interface. IEEE, pp. 1–3.
Koesdwiady, A., Bedawi, S.M., Ou, C., Karray, F., 2017a. End-to-end deep learning for driver distraction recognition. In: International Conference Image Analysis and Recognition. Springer, pp. 11–18.
Koesdwiady, A., Soua, R., Karray, F., Kamel, M.S., 2017b. Recent trends in driver safety monitoring systems: State of the art and challenges. IEEE Trans. Veh. Technol. 66 (6), 4550–4563.
Kong, W., Zhou, Z., Jiang, B., Babiloni, F., Borghini, G., 2017. Assessment of driving fatigue based on intra/inter-region phase synchronization. Neurocomputing 219, 474–482.
Korbar, B., Tran, D., Torresani, L., 2018. Cooperative learning of audio and video models from self-supervised synchronization. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. NIPS '18, Curran Associates Inc., Red Hook, NY, USA, pp. 7774–7785.
Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., Torralba, A., 2016. Eye tracking for everyone. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2176–2184.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25, 1097–1105.
Lal, S.K., Craig, A., 2001a. A critical review of the psychophysiology of driver fatigue. Biol. Psychol. 55 (3), 173–194.
Lal, S.K., Craig, A., 2001b. Electroencephalography activity associated with driver fatigue: Implications for a fatigue countermeasure device. J. Psychophysiol. 15 (3), 183.
Lal, S.K., Craig, A., Boord, P., Kirkup, L., Nguyen, H., 2003. Development of an algorithm for an EEG-baseddriver fatigue countermeasure. J. Saf. Res. 34 (3), 321–328.
Lee, S., Kim, J., Shin Yoon, J., Shin, S., Bailo, O., Kim, N., Lee, T.-H., Seok Hong, H., Han, S.-H., So Kweon, I., 2017. VPGNet: Vanishing point guided network for lane and road marking detection and recognition. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 1947–1955.
Li, Z., Chen, L., Peng, J., Wu, Y., 2017. Automatic detection of driver fatigue using driving operation information for transportation safety. Sensors 17 (6), 1212.

Li, X., Hong, L., Wang, J.-c., Liu, X., 2019a. Fatigue driving detection model based on multi-feature fusion and semi-supervised active learning. IET Intell. Transp. Syst. 13 (9), 1401–1409.
Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J., 2021. A survey of convolutional neural networks: Analysis, applications, and prospects. IEEE Trans. Neural Netw. Learn. Syst. 1–21.
Li, X., Seignez, E., Lambert, A., Loonis, P., 2014. Real-time driver drowsiness estimation by multi-source information fusion with Dempster–Shafer theory. Trans. Inst. Meas. Control 36 (7), 906–915.
Li, Z., Yang, Q., Chen, S., Zhou, W., Chen, L., Song, L., 2019b. A fuzzy recurrent neural network for driver fatigue detection based on steering-wheel angle sensor data. Int. J. Distrib. Sens. Netw. 15 (9), 1550147719872452.
Li, H., Zheng, W.-L., Lu, B.-L., 2018. Multimodal vigilance estimation with adversarial domain adaptation networks. In: International Joint Conference on Neural Networks. IEEE, pp. 1–6.
Lim, S., Yang, J.H., 2016. Driver state estimation by convolutional neural network using multimodal sensor data. Electron. Lett. 52 (17), 1495–1497.
Liu, F., Li, X., Lv, T., Xu, F., 2019a. A review of driver fatigue detection: Progress and prospect. In: 2019 IEEE International Conference on Consumer Electronics. IEEE, pp. 1–6.
Liu, X., Liang, W., Wang, Y., Li, S., Pei, M., 2016. 3D head pose estimation with convolutional neural network trained on synthetic images. In: IEEE International Conference on Image Processing. IEEE, pp. 1289–1293.
Liu, Y.-T., Lin, Y.-Y., Wu, S.-L., Chuang, C.-H., Lin, C.-T., 2015. Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. IEEE Trans. Neural Netw. Learn. Syst. 27 (2), 347–360.
Liu, W., Qian, J., Yao, Z., Jiao, X., Pan, J., 2019b. Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection. Future Internet 11 (5), 115.
Liu, F., Tang, Z., Tang, J., 2013. WLBP: Weber local binary pattern for local image description. Neurocomputing 120, 325–335, Image Feature Detection and Description.
Liu, S., Yu, L., Hou, M., 2019c. An efficient method for driver fatigue state detection based on deep learning. In: The 2nd International Conference on Safety Produce Informatization. IEEE, pp. 172–176.
Lopez, M.B., del Blanco, C.R., Garcia, N., 2017. Detecting exercise-induced fatigue using thermal imaging and deep learning. In: The 7th International Conference on Image Processing Theory, Tools and Applications. IEEE, pp. 1–6.
Lu, J., Zheng, X., Tang, L., Zhang, T., Sheng, Q.Z., Wang, C., Jin, J., Yu, S., Zhou, W., 2021. Can steering wheel detect your driving fatigue? IEEE Trans. Veh. Technol. 70 (6), 5537–5550.
Luo, H., Qiu, T., Liu, C., Huang, P., 2019. Research on fatigue driving detection using forehead EEG based on adaptive multi-scale entropy. Biomed. Signal Process. Control 51, 50–58.
Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J., 1998. Coding facial expressions with Gabor wavelets. In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. pp. 200–205.
Lyu, J., Yuan, Z., Chen, D., 2018. Long-term multi-granularity deep framework for driver drowsiness detection. arXiv preprint arXiv:1801.02325.
Ma, X., Chau, L.-P., Yap, K.-H., Ping, G., 2019a. Convolutional three-stream network fusion for driver fatigue detection from infrared videos. In: IEEE International Symposium on Circuits and Systems. IEEE, pp. 1–5.
Ma, Y., Chen, B., Li, R., Wang, C., Wang, J., She, Q., Luo, Z., Zhang, Y., 2019b. Driving fatigue detection from EEG using a modified PCANet method. Comput. Intell. Neurosci. 2019.
Marinello, F., Pezzuolo, A., Gasparini, F., Sartori, L., 2014. Ergonomics analyses through motion capture in a vehicle cabin by means of Kinect sensor. In: International Conference of Agricultural Engineering, Vol. 253. Citeseer, pp. 1–5.
Martínez, A., 1998. The AR face database.
Mashko, A., 2015. Review of approaches to the problem of driver fatigue and drowsiness. In: 2015 Smart Cities Symposium Prague. IEEE, pp. 1–5.
Matsushita, Y., Miura, J., 2011. On-line road boundary modeling with multiple sensory features, flexible road model, and particle filter. Robot. Auton. Syst. 59 (5), 274–284.
Meireles, T., Dantas, F., 2019. A low-cost prototype for driver fatigue detection. Multimodal Technol. Interact. 3 (1), 5.
Meng, F., Li, S., Cao, L., Li, M., Peng, Q., Wang, C., Zhang, W., 2015. Driving fatigue in professional drivers: a survey of truck and taxi drivers. Traffic Inj. Prev. 16 (5), 474–483.
Mobileye Company, Mobileye Company, [EB/OL], 2020. https://www.mobileye.com/ Accessed December 1, 2020.
Murphy-Chutorian, E., Trivedi, M.M., 2010. Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness. IEEE Trans. Intell. Transp. Syst. 11 (2), 300–311.
Naaz, A., Singh, M.S., 2014. Feature extraction and analysis of ECG signal for cardiac abnormalities: A review. Int. J. Eng. Res. Technol. 3 (11), 23–30.
Naeije, M., Zorn, H., 1982. Relation between EMG power spectrum shifts and muscle fibre action potential conduction velocity changes during local muscular fatigue in man. Eur. J. Appl. Physiol. Occup. Physiol. 50 (1), 23–33.
Nanjing Yuanqu Technology Co., Ltd. [EB/OL]. 2020. http://www.fardriver.com/. Accessed December 1, 2020.
Noguchi, Y., Nopsuwanchai, R., Ohsuga, M., Kamakura, Y., 2007. Classification of blink waveforms towards the assessment of driver's arousal level-an approach for hmm based classification from blinking video sequence. In: International Conference on Engineering Psychology and Cognitive Ergonomics. Springer, pp. 779–786.
Němcová, A., Svozilová, V., Bucsuházy, K., Smíšek, R., Mézl, M., Hesko, B., Belák, M., Bilík, M., Maxera, P., Seitl, M., Dominik, T., Semela, M., Šucha, M., Kolář, R., 2020. Multimodal features for detection of driver stress and fatigue: Review. IEEE Trans. Intell. Transp. Syst. 1–20.
Němcová, A., Svozilová, V., Bucsuházy, K., Smíšek, R., Mézl, M., Hesko, B., Belák, M., Bilík, M., Maxera, P., Seitl, M., Dominik, T., Semela, M., Šucha, M., Kolář, R., 2021. Multimodal features for detection of driver stress and fatigue: Review. IEEE Trans. Intell. Transp. Syst. 22 (6), 3214–3233.
Ohsuga, M., Kamakura, Y., Inoue, Y., Noguchi, Y., Nopsuwanchai, R., 2007. Classification of blink waveforms toward the assessment of driver's arousal levels-an EOG approach and the correlation with physiological measures. In: International Conference on Engineering Psychology and Cognitive Ergonomics. Springer pp. 787–795.
Ojala, T., Pietikainen, M., Maenpaa, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. 24 (7), 971–987.
Omerustaoglu, F., Sakar, C.O., Kar, G., 2020. Distracted driver detection by combining in-vehicle and image data using deep learning. Appl. Soft Comput. 96, 106657.
Owen, N., King, H., Lamb, M., 2015. Literature review of race driver fatigue measurement in endurance motorsport. Procedia Eng. 112, 344–348.
Panwar, S., Rad, P., Quarles, J., Golob, E., Huang, Y., 2019. A semi-supervised wasserstein generative adversarial network for classifying driving fatigue from EEG signals. In: IEEE International Conference on Systems, Man and Cybernetics. IEEE, pp. 3943–3948.
Park, S., Pan, F., Kang, S., Yoo, C.D., 2016. Driver drowsiness detection system based on feature representation learning using various deep networks. In: Asian Conference on Computer Vision. Springer, pp. 154–164.
Patrick, M., Asano, Y.M., Fong, R., Henriques, J.F., Zweig, G., Vedaldi, A., 2020. Multi-modal self-supervision from generalized data transformations. CoRR abs/2003.04298.
Pauly, L., Sankar, D., 2015. Detection of drowsiness based on HOG features and SVM classifiers. In: 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks. IEEE, pp. 181–186.
1998. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance.
Petrofsky, J.S., Lind, A.R., 1980. The influence of temperature on the amplitude and frequency components of the EMG during brief and sustained isometric contractions. Eur. J. Appl. Physiol. Occup. Physiol. 44 (2), 189–200.
Pioneer Corporation, [EB/OL]. 2020. https://global.pioneer/en/. Accessed December 1, 2020.
Poorna, S., Deep, A., Hariharan, K., Jain, R.R., Krishnan, S., 2021. A transfer learning approach for drowsiness detection from EEG signals. In: Innovations in Computational Intelligence and Computer Vision. Springer, pp. 369–375.
Qiao, Y., Zeng, K., Xu, L., Yin, X., 2016. A smartphone-based driver fatigue detection using fusion of multiple real-time facial features. In: The 13th IEEE Annual Consumer Communications & Networking Conference. IEEE, pp. 230–235.
Radford, A., Metz, L., Chintala, S., 2016. Unsupervised representation learning with deep convolutional generative adversarial networks. In: Bengio, Y., LeCun, Y. (Eds.), 4th International Conference on Learning Representations.
Raez, M.B., Hussain, M.S., Mohd-Yasin, F., 2006. Techniques of EMG signal analysis: detection, processing, classification and applications. Biol. Proced. Online 8, 11–35.
Ramzan, M., Khan, H.U., Awan, S.M., Ismail, A., Ilyas, M., Mahmood, A., 2019. A survey on state-of-the-art drowsiness detection techniques. IEEE Access 7, 61904–61919.
Reddy, B., Kim, Y.-H., Yun, S., Seo, C., Jang, J., 2017. Real-time driver drowsiness detection for embedded system using model compression of deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 121–128.
Rogado, E., Garcia, J.L., Barea, R., Bergasa, L.M., López, E., 2009. Driver fatigue detection system. In: IEEE International Conference on Robotics and Biomimetics. IEEE, pp. 1105–1110.
Rongben, W., Lie, G., Bingliang, T., Lisheng, J., 2004. Monitoring mouth movement for driver fatigue or distraction with one camera. In: The 7th International IEEE Conference on Intelligent Transportation Systems. IEEE, pp. 314–319.
Rundo, F., Rinella, S., Massimino, S., Coco, M., Fallica, G., Parenti, R., Conoci, S., Perciavalle, V., 2019. An innovative deep learning algorithm for drowsiness detection from EEG signal. Computation 7 (1), 13.
SafeTrac Solutions, [EB/OL], 2020. https://www.safetracsolutions.com/products. Accessed December 1, 2020.
Sahayadhas, A., Sundaraj, K., Murugappan, M., 2012. Detecting driver drowsiness based on sensors: a review. Sensors 12 (12), 16937–16953.
San, P.P., Ling, S.H., Chai, R., Tran, Y., Craig, A., Nguyen, H., 2016. EEG-based driver fatigue detection using hybrid deep generic model. In: The 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBC, IEEE, pp. 800–803.

Sandberg, D., Wahde, M., 2008. Particle swarm optimization of feedforward neural networks for the detection of drowsy driving. In: IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). IEEE, pp. 788–793.

Sanjaya, K.H., Lee, S., Katsuura, T., 2016. Review on the application of physiological and biomechanical measurement methods in driving fatigue detection. J. Mechatron. Electr. Power Veh. Technol. 7 (1), 35–48.

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. Springer pp. 189–200.

Shafer, G., 1976. A Mathematical Theory of Evidence, Vol. 1. PrincetonUniversity Press.

Shahverdy, M., Fathy, M., Berangi, R., Sabokrou, M., 2020. Driver behavior detection and classification using deep convolutional neural networks. Expert Syst. Appl. 149, 113240.

Shen, K.-Q., Li, X.-P., Ong, C.-J., Shao, S.-Y., Wilder-Smith, E.P., 2008. EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate. Clin. Neurophysiol. 119 (7), 1524–1533.

Shi, S.-Y., Tang, W.-Z., Wang, Y.-Y., 2017. A review on fatigue driving detection. In: ITM Web of Conferences, Vol. 12. EDP Sciences, p. 01019.

Shih, T.-H., Hsu, C.-T., 2016. MSTN: Multistage spatial-temporal network for driver drowsiness detection. In: Asian Conference on Computer Vision. Springer pp. 146–153.

Shin, J., Kim, S., Yoon, T., Joo, C., Jung, H.-I., 2019. Smart Fatigue Phone: Real-time estimation of driver fatigue using smartphone-based cortisol detection. Biosens. Bioelectron. 136, 106–111.

Sikander, G., Anwar, S., 2018. Driver fatigue detection systems: A review. IEEE Trans. Intell. Transp. Syst. 20 (6), 2339–2352.

Sikander, G., Anwar, S., 2021. A novel machine vision-based 3D facial action unit identification for fatigue detection. IEEE Trans. Intell. Transp. Syst. 22 (5), 2730–2740.

SP, R.M., Reshma, D., et al., 2021. A real-time fatigue detection system using multi-task cascaded CNN model. In: 2021 10th IEEE International Conference on Communication Systems and Network Technologies. CSNT, IEEE, pp. 674–679.

Stork, M., Skála, J., Weissar, P., Holota, R., Kubík, Z., 2015. Various approaches to driver fatigue detection: A review. In: 2015 International Conference on Applied Electronics. IEEE, pp. 239–244.

Streiffer, C., Raghavendra, R., Benson, T., Srivatsa, M., 2017. Darnet: a deep learning solution for distracted driving detection. In: The 18th ACM/IFIP/USENIX Middleware Conference: Industrial Track. pp. 22–28.

Suzuki, M., Yamamoto, N., Yamamoto, O., Nakano, T., Yamamoto, S., 2006. Measurement of driver's consciousness by image processing-a method for presuming driver's drowsiness by eye-blinks coping with individual differences. In: IEEE International Conference on Systems, Man and Cybernetics. 4, IEEE, pp. 2891–2896.

Takei, Y., Furukawa, Y., 2005. Estimate of driver's fatigue through steering motion. In: IEEE International Conference on Systems, Man and Cybernetics, Vol. 2. IEEE pp. 1765–1770.

Tango, F., Botta, M., 2013. Real-time detection system of driver distraction using machine learning. IEEE Trans. Intell. Transp. Syst. 14 (2), 894–905.

Tefft, B.C., 2012. Prevalence of motor vehicle crashes involving drowsy drivers, United States, 1999–2008. Accid. Anal. Prev. 45, 180–186.

Tran, Y., Wijesuryia, N., Thuraisingham, R.A., Craig, A., Nguyen, H.T., 2008. Increase in regularity and decrease in variability seen in electroencephalography (EEG) signals from alert to fatigue during a driving simulated task. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE pp. 1096–1099.

2020. System for effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation. [EB/OL], https://trimis.ec.europa.eu/entityprint/node/4462.

Tsai, Y.-C., Lai, P.-W., Huang, P.-W., Lin, T.-M., Wu, B.-F., 2020. Vision-based instant measurement system for driver fatigue monitoring. IEEE Access 8, 67342–67353.

Tuncer, T., Dogan, S., Ertam, F., Subasi, A., 2020. A dynamic center and multi threshold point based stable feature extraction network for driver fatigue detection utilizing EEG signals. Cogn. Neurodyn. 1–15.

Utomo, D., Yang, T.-H., Thanh, D.T., Hsiung, P.-A., 2019. Driver fatigue prediction using different sensor data with deep learning. In: IEEE International Conference on Industrial Cyber Physical Systems. IEEE, pp. 242–247.

Venturelli, M., Borghi, G., Vezzani, R., Cucchiara, R., 2017. From depth data to head pose estimation: A siamese approach. In: Imai, F.H., Trémeau, A., Braz, J. (Eds.), Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. SciTePress, pp. 194–201. http://dx.doi.org/10.5220/0006104501940201.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A., 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. J. Mach. Learn. Res. 11, 3371–3408.

Wang, J., Chai, W., Venkatachalapathy, A., Tan, K.L., Haghighat, A., Velipasalar, S., Adu-Gyamfi, Y., Sharma, A., 2021. A survey on driver behavior analysis from in-vehicle cameras. IEEE Trans. Intell. Transp. Syst..

Wang, M., Deng, W., 2018. Deep visual domain adaptation: A survey. Neurocomputing 312, 135–153.

Wang, P., Ji, Q., 2005. Learning discriminant features for multi-view face and eye detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1. IEEE, pp. 373–379.

Wang, Q., Yang, J., Ren, M., Zheng, Y., 2006. Driver fatigue detection: a survey. In: 2006 6th World Congress on Intelligent Control and Automation, Vol. 2. IEEE pp. 8587–8591.

Wen, T., Zhang, Z., 2018. Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals. IEEE Access 6, 25399–25410.

Weng, C.-H., Lai, Y.-H., Lai, S.-H., 2016. Driver drowsiness detection via a hierarchical temporal deep belief network. In: Asian Conference on Computer Vision. Springer, pp. 117–133.

Wierwille, W.W., Wreggit, S., Kirn, C., Ellsworth, L., Fairbanks, R., 1994. Research on Vehicle-Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness. Final Report.

Williamson, A., Chamberlain, T., 2005. Review of On-Road Driver Fatigue Monitoring Devices. NSW Injury Risk Management Research Centre, University of New South Wales.

Wongphanngam, J., Pumrin, S., 2016. Fatigue warning system for driver nodding off using depth image from Kinect. In: The 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). IEEE, pp. 1–6.

Wu, E.Q., Peng, X., Zhang, C.Z., Lin, J., Sheng, R.S., 2019. Pilots' fatigue status recognition using deep contractive autoencoder network. IEEE Trans. Instrum. Meas. 68 (10), 3907–3919.

Wu, W., Wu, Q.M.J., Sun, W., Yang, Y., Yuan, X., Zheng, W.-L., Lu, B.-L., 2021. A regression method with subnetwork neurons for vigilance estimation using EOG and EEG. IEEE Trans. Cogn. Dev. Syst. 13 (1), 209–222.

Wulsin, D., Gupta, J., Mani, R., Blanco, J., Litt, B., 2011. Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement. J. Neural Eng. 8 (3), 036015.

Xiao, Z., Hu, Z., Geng, L., Zhang, F., Wu, J., Li, Y., 2019. Fatigue driving recognition network: Fatigue driving recognition via convolutional neural network and long short-term memory units. IET Intell. Transp. Syst. 13 (9), 1410–1416.

Xie, Y., Chen, K., Murphey, Y.L., 2018. Real-time and robust driver yawning detection with deep neural networks. In: IEEE Symposium Series on Computational Intelligence. IEEE, pp. 532–538.

Xie, Y., Li, F., Wu, Y., Yang, S., Wang, Y., 2019. D3-Guard: Acoustic-based drowsy driving detection using smartphones. In: IEEE Conference on Computer Communications. IEEE, pp. 1225–1233.

Xing, Y., Lv, C., Cao, D., 2020. Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles. IEEE Trans. Veh. Technol. 69 (2), 1341–1352.

Yang, G., Lin, Y., Bhattacharya, P., 2005. A driver fatigue recognition model using fusion of multiple features. In: IEEE International Conference on Systems, Man and Cybernetics, Vol. 2. IEEE, pp. 1777–1784.

Yang, G., Lin, Y., Bhattacharya, P., 2010. A driver fatigue recognition model based on information fusion and dynamic Bayesian network. Inform. Sci. 180 (10), 1942–1954.

Yang, H., Liu, L., Min, W., Yang, X., Xiong, X., 2021. Driver yawning detection based on subtle facial action recognition. IEEE Trans. Multimed. 23, 572–583.

Yang, C., Wang, X., Mao, S., 2020. Unsupervised drowsy driving detection with RFID. IEEE Trans. Veh. Technol. 69 (8), 8151–8163.

Ye, W., Zhao, X., 2018. Processing and filtrating of driver fatigue characteristic parameters based on rough set. In: AIP Conference Proceedings, Vol. 1967. AIP Publishing LLC, 020001.

Yu, J., Park, S., Lee, S., Jeon, M., 2016. Representation learning, scene understanding, and feature fusion for drowsiness detection. In: Asian Conference on Computer Vision. Springer, pp. 165–177.

Zhang, Y., Chen, Y., Gao, C., 2021. Deep unsupervised multi-modal fusion network for detecting driver distraction. Neurocomputing 421, 26–38.

Zhang, G., Etemad, A., 2019. Capsule attention for multimodal EEG and EOG spatiotemporal representation learning with application to driver vigilance estimation. arXiv preprint arXiv:1912.07812.

Zhang, L., Liu, F., Tang, J., 2015a. Real-time system for driver fatigue detection by RGB-D camera. ACM Trans. Intell. Syst. Technol. (TIST) 6 (2), 1–17.

Zhang, C., Lu, X., Huang, Z., Xia, S., Fu, C., 2019. A driver fatigue recognition algorithm based on spatio-temporal feature sequence. In: The12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. IEEE pp. 1–6.

Zhang, W., Murphey, Y.L., Wang, T., Xu, Q., 2015b. Driver yawning detection based on deep convolutional neural learning and robust nose tracking. In: International Joint Conference on Neural Networks. IEEE, pp. 1–8.

Zhang, F., Su, J., Geng, L., Xiao, Z., 2017. Driver fatigue detection based on eye state recognition. In: International Conference on Machine Vision and Information Technology. IEEE, pp. 105–110.

Zhang, X., Sugano, Y., Fritz, M., Bulling, A., 2015c. Appearance-based gaze estimation in the wild. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 4511–4520.

Zhang, C., Wang, H., Fu, R., 2013. Automated detection of driver fatigue based on entropy and complexity measures. IEEE Trans. Intell. Transp. Syst. 15 (1), 168–177.

Zhang, W., Wang, F., Wu, S., Xu, Z., Ping, J., Jiang, Y., 2020. Partial directed coherence based graph convolutional neural networks for driving fatigue detection. Rev. Sci. Instrum. 91 (7), 074713.

Zhang, N., Zheng, W.-L., Liu, W., Lu, B.-L., 2016. Continuous vigilance estimation using lstm neural networks. In: International Conference on Neural Information Processing. Springer, pp. 530–537.

Zhao, S., Chen, G., Hua, M., Zong, C., 2019. An identification algorithm of driver steering characteristics based on backpropagation neural network. Proc. Inst. Mech. Eng. D 233 (9), 2333–2342.

Zhao, L., Wang, Z., Wang, X., Liu, Q., 2017. Driver drowsiness detection using facial dynamic fusion information and a DBN. IET Intell. Transp. Syst. 12 (2), 127–133.

Zhao, L., Wang, Z., Zhang, G., Qi, Y., Wang, X., 2018. Eye state recognition based on deep integrated neural network and transfer learning. Multimedia Tools Appl. 77 (15), 19415–19438.

Zhong, Y.-J., Du, L.-P., Zhang, K., Sun, X.-H., 2007. Localized energy study for analyzing driver fatigue state based on wavelet analysis. In: International Conference on Wavelet Analysis and Pattern Recognition, Vol. 4. IEEE, pp. 1843–1846.

Zhu, W., Yang, H., Jin, Y., Liu, B., 2017. A method for recognizing fatigue driving based on Dempster–Shafer theory and fuzzy neural network. Math. Probl. Eng. 2017, 6191035.

Zhu, X., Zheng, W.-L., Lu, B.-L., Chen, X., Chen, S., Wang, C., 2014. EOG-based drowsiness detection using convolutional neural networks. In: International Joint Conference on Neural Networks. IEEE, pp. 128–134.

Zou, Q., Jiang, H., Dai, Q., Yue, Y., Chen, L., Wang, Q., 2019. Robust lane detection from continuous driving scenes using deep neural networks. IEEE Trans. Veh. Technol. 69 (1), 41–54.

Zuraida, R., Iridiastadi, H., Sutalaksana, I.Z., Suprijanto, S., 2019. An analysis of EEG changes during prolonged simulated driving for the assessment of driver fatigue. J. Eng. Technol. Sci. 51 (2), 290.