Theory and Practice of Science and Technology
2025, VOL. 6, NO. 7, 79-81
DOI: 10.47297/taposatWSP2633-456916.20250607

Analysis of Research for Fatigue Driving Detection in Complex Scenarios

Sirui Min*

Chang'an Dublin International Transportation College, Chang'an University, Xi'an, Shaanxi, 710000, China

ABSTRACT
Fatigued driving is a major threat to road traffic safety, and its identification and warning become more challenging in complex environments such as at night, in adverse weather conditions, in mountainous areas, and in urban congestion. This paper examines how complex environmental factors influence fatigue detection and the current research in detection using multimodal perception technology. This includes methods based on physiological signals, as well as techniques for extracting facial landmarks and blink parameters under low-light conditions or facial occlusion. This paper further explores multimodal fusion methods and their application in improving the detection accuracy of environmental factors. Finally, addressing the shortcomings of current research in algorithm robustness, scene generalization capability, and real-road validation, the paper summarizes the challenges faced and outlines future directions aiming to provide references for in-depth research and intelligent prevention and control practices regarding fatigue driving in complex environments.

KEYWORDS
Complex scenes; Fatigue detection; Traffic safety; Multimodal fusion

1 Introduction

Road traffic safety is a major global public safety issue. When drivers are fatigued, their perception, judgment, and reaction speed are significantly impaired, making them extremely prone to traffic accidents. In reality, driving often takes place in complex and changing environments, such as low light conditions when driving at night, reduced visibility due to heavy rain and smog, sharp bends and slopes on mountain roads, and urban congestion. This significantly increases the cognitive load and operational difficulty of driving. In recent years, the rapid development of multimodal perception technology has provided new insights into fatigue detection. Researchers have combined environmental information to achieve a more comprehensive analysis of driver states in complex scenarios. However, algorithms still face challenges in noise suppression, scene adaptation, and real-time embedded deployment. This paper systematically reviews the impact of complex environmental factors on fatigue performance, analyzes the various detection technologies, and finally discusses the prospects for algorithm robustness and scene generalization, providing a reference for research on fatigue driving in complex environments.

2 Challenges in detecting fatigue driving under complex conditions

The complex and ever-changing road environment poses significant interference and challenges to current mainstream fatigue driving detection technologies. First, environmental uncertainty significantly increases the difficulty of detection. Factors such as changes in lighting, adverse weather conditions, and obstructions interfere with the collection of facial features and physiological signals, reducing the robustness and stability of the system. Second, there are significant differences in fatigue manifestations among individual drivers, including blink frequency, facial movements, and behavioral habits, making it difficult for a unified model to accurately adapt to all users. Additionally, complex scenarios often require the integration of multimodal data to improve detection accuracy, but multi-source data poses high technical barriers in terms of time synchronization, feature alignment, and fusion algorithm design. Meanwhile, the acquisition and annotation of high-quality fatigue data, especially in real-world complex environments, is extremely costly, severely limiting the training and deployment of deep learning models. Finally, to achieve practical deployment, the detection system must not only have high recognition accuracy but also meet real-time requirements. However, on resource-constrained in-vehicle computing platforms, balancing model complexity and processing speed remains an urgent issue to be addressed. As such, fatigue driving detection under complex conditions is both a technical challenge and a key direction for future research.

3 Various fatigue driving detection methods

3.1 Based on physiological signals

Li Gang's team has developed a technology based on electroencephalography (EEG), which is considered one of the

*** Corresponding Author:** Sirui Min, sirui.min@ucdconnect.ie.

core methods for fatigue detection because it can directly capture neural activity signals that reflect the brain's fatigue state. This technology can identify a driver's early micro-sleep state within 10-30 seconds by analyzing characteristics such as the power ratio of alpha waves and theta waves. However, vehicle vibrations and driver head movements can easily distort EEG signals, while electrocardiogram (ECG) and electrooculogram (EOG) signals are often interfered with by electromyographic noise.

To address these signal quality issues, the study employed wavelet packet transform (WPT) to process the non-stationary characteristics of EEG signals and perform noise reduction. Additionally, motion sensors such as gyroscopes were introduced to real-time identify and correct artifacts caused by movement. At the hardware level, optimized designs such as O1/O2 occipital bipolar leads and dry electrodes were adopted to enhance the stability of signal acquisition.

On the other hand, there are significant individual differences in EEG baselines among different drivers, which limit the model's generalization ability. To address the issue of individual differences, the study gradually introduced transfer learning strategies to enhance the model's adaptability to different users. Additionally, personalized calibration methods were combined, setting dynamic detection thresholds based on individual EEG features in the awake state, thereby effectively improving the model's recognition accuracy.

3.2 Facial recognition-based

In recent years, research on fatigue detection based on driver facial features has become mainstream. However, existing methods still face severe challenges in complex driving scenarios such as drastic changes in lighting, facial obstruction, or large-angle head deviation.

Chen Shili and others proposed a pure vision-based detection scheme using image processing. This method rapidly locates the eye region using image projection and morphological operations, and identifies fatigue states by calculating the width-to-height ratio of the eyes (a ratio greater than 2.5 indicates closed eyes) and the proportion of closed eye time (a PERCLOS value greater than 1/3 indicates fatigue). However, this scheme has weak interference resistance. Under conditions such as strong light exposure, the driver wearing sunglasses, or head rotation, eye localization often fails, leading to a significant increase in false positive rates.

To address dynamic occlusion and large-angle head rotation issues, Zhichao Sun et al. proposed the FFF-CNN model. This model adopts a three-stream network architecture. When the driver's hands occlude the eyes or the head turns more than 30°, the accuracy of traditional single-stream networks drops below 90%. In contrast, FFF-CNN enhances key region features through the feature attention module (FAM) and corrects inconsistencies in binocular features using the stream interaction module (SIM), improving detection accuracy to 94.89%. Notably, in scenarios where both eyes are in different states, such as one eye being closed, FFF-CNN reduces the false positive rate by 83%.

To address the challenges of strong light interference and sunglasses obstruction, Zhan Hongmei et al. designed a collaborative solution combining a 940nm near-infrared active light source with a bandpass filter. This scheme effectively eliminates visible light interference in glare or nighttime environments, improving image quality stability by 40%; for drivers wearing sunglasses, near-infrared light can penetrate the lenses for imaging, addressing the false negative rate issue caused by lens reflections in traditional methods (where the false negative rate exceeded 60%).

Hu Xizhi et al. constructed a cascaded framework integrating SSD (ResNet 50), the CamShift tracker, and a Kalman filter. The framework first uses SSD to extract high-precision face bounding boxes; when occlusion or lighting changes occur, CamShift combines the Bhattacharyya coefficient for tracking, supplemented by Kalman filtering to predict the target's position; Finally, a cascaded regression tree is used to locate eye and mouth feature points to extract facial features. This design effectively addresses image blurring caused by vehicle vibrations and temporary target loss. Experiments show that on mountainous, bumpy roads, the Kalman filter's prediction function reduces the CamShift tracking loss rate from 21% to 3%. After the driver briefly obstructs the face by looking down, the system can resume detection within 0.1 seconds after the target reappears. This dynamic cascaded framework achieves an average accuracy of 92.2% in real-vehicle environments, with a 50% reduction in false alarm rate, significantly enhancing the system's continuous detection capability in complex road conditions.

3.3 Multimodal fusion

The advantage of multimodal fusion methods lies in their ability to integrate multiple information sources, including visual, electroencephalographic (EEG), physiological signals, and driving behavior. This not only effectively compensates for the shortcomings of single-modal approaches in complex situations such as sudden changes in lighting, obstruction, posture deviation, or signal loss, but also enhances the ability to characterize and distinguish fatigue states. The core idea is that when one information source fails, other modalities can still provide effective judgment criteria.

Research by Gao Qihuang et al. proposed a fusion process: first, accurately extract the driver's facial and physiological features in complex driving environments; then use Restricted Boltzmann Machines (RBM) for deep information mining, and through a Multi-Loss Reconstruction (MLR) module, selectively fuse the deep fatigue features of each modality while removing redundant information; finally, based on a Bidirectional Long Short-Term Memory Network (Bi-LSTM), model the fused temporal features to achieve efficient identification of the driver's fatigue state.

To enhance detection robustness in complex environments, Sun Jian et al. proposed a detection method that integrates machine vision with vehicle trajectory information. In terms of vision, they based their approach on the ResNet50 network architecture and optimized the loss function, improving the face keypoint detection algorithm to accurately locate 98 keypoints on the face under different lighting conditions and head postures. Using these keypoints, three types of visual features were extracted: eye opening degree and closed eye ratio, mouth yawning degree, and head nodding frequency and pitch angle. In terms of vehicle behavior, the method combines Beidou high-precision positioning technology (RTK) to calculate the slope of longitudinal displacement fluctuations (SLDF), utilizing the characteristic that vehicle trajectory fluctuations tend to flatten during fatigued driving as an auxiliary judgment criterion. Finally, an improved support vector machine (SVM) classification algorithm is used to fuse the aforementioned four types of features for comprehensive judgment. However, this method relies on high-precision positioning equipment, which is costly and may fail in extreme scenarios such as strong light inside tunnels. Additionally, the current algorithm's overall processing time is approximately 3 seconds, and its real-time performance requires further improvement.

Cheng Wenxin et al. proposed a multi-modal feature fusion model named nsNMF-PCNN-GRU-MSA and combined it with EEG analysis of physiological signals. The model first uses the non-smooth non-negative matrix factorization (nsNMF) algorithm to calculate the weights of different electrode channels, highlighting activity in the occipital and parietal regions during fatigue. Subsequently, Gram angle field imaging technology is used to convert the one-dimensional EEG time series into a two-dimensional image. The model further employs a parallel structure to fuse spatial and temporal features and introduces a multi-head self-attention mechanism for final classification. Since EEG signals directly reflect the brain's physiological state, this method effectively avoids interference from external environmental factors such as lighting changes and facial occlusion. However, while the use of multi-channel electrode caps can obtain rich EEG information, their poor wearing comfort may affect user experience in long-term driving scenarios.

4 Conclusion

The dynamic characteristics of complex road environments severely interfere with almost all fatigue detection technologies, so research into the detection in complex driving environments is getting more attention. With the development of artificial intelligence and sensor technologies, fatigue detection methods are becoming more diverse, mainly including those based on facial features, physiological signals, driving behavior, and multimodal fusion methods. Facial feature detection has the advantages of being non-contact and real-time, but it's easily affected by factors like lighting and obstruction. Physiological signal detection, while highly accurate, requires the use of devices, raising privacy and comfort concerns; driving behavior detection does not require additional devices but has limited adaptability to different driving styles. Additionally, multi-modal fusion methods enhance detection robustness and accuracy by combining multiple detection approaches. However, current research still faces challenges such as significant individual variability, strong environmental interference, and inconsistent data annotation. Future research should focus on developing more lightweight, privacy-protected detection methods, constructing intelligent fatigue monitoring systems adaptable to complex scenarios, improving the efficiency and accuracy of multimodal data fusion, and strengthening testing and validation in real-world driving scenarios.

References

[1] Cheng, W.X., Yan, G.H., Chang, W.W. et al. Based on channel weighted multimodal feature fusion for EEG fatigue driving detection. Journal of Zhejiang University (Engineering Science), P. 1–10. 2025.

[2] Sun, J., Tang, X., Xu, Y.N. et al. Detection method of highway fatigue driving behavior by fusing machine vision and high-precision positioning. Transportation Research, 2023, 9(6), 78–87+118.

[3] Sun, Z.C., Miao, Y.N., Jeon, J.Y. et al. Facial feature fusion convolutional neural network for driver fatigue detection. Engineering Applications of Artificial Intelligence, P. 126(PC). 2023.

[4] Gao, Q. H., Xie, K., He, Z. F. et al. Fatigue driving detection based on multimodal feature fusion in complex environment. Electronic Measurement Technology, 2023, 46(6), 106–115.

[5] Li, G. & Chung, W.Y. Electroencephalogram-Based Approaches for Driver Drowsiness Detection and Management: A Review. Sensors, 2022, 22 (3), 1100–1100.

[6] Hu, X.Z. & Huang, B.Y. Fatigue driving detection method based on facial feature analysis. Science Technology and Engineering, 2021, 21(4), 1629–1636.

[7] Zhan, H.M. Anti fatigue driving system based on facial feature recognition technology in complex driving environment. Electronic Test, 2016, (17), 84–85.

[8] Chen, S.L., Lei, L. & Zhao, Y.X. Research on fatigue driving detection technology based on behavioral characteristics. Journal of Chengdu University (Natural Science Edition), 2014, 33(1), 37–40.