Check for updates

Comparative Analysis of Deep Learning Methods Using Multiple Modal Data for Driver Fatigue Identification

Yian Wang

North Broward Preparatory School, FL, 33073, United States of America
yi-an_wang@mynbps.org

Abstract. Driver fatigue is a critical safety concern that significantly increases the risk of accidents on roads. By analyzing patterns and behaviors through sophisticated algorithms, deep learning can predict fatigue states and alert drivers, thereby enhancing safety. Its predictive capabilities can also inform the development of systems that promote driver safety and reduce the likelihood of fatigue-related incidents. This paper examines various machine learning-based approaches to detecting driver fatigue, focusing primarily on image-based and physiological signal-based fatigue identification. Image-based techniques utilize facial recognition and behavioral analysis to detect fatigue signs, such as frequent blinking and yawning, while physiological methods analyze data from sensors that measure heart rate variability, brain waves, and other bodily signals indicative of fatigue. The review highlights the strengths and limitations of each method, emphasizing the potential of integrating these approaches in multimodal systems. Challenges such as variable environmental conditions, system reliability, and data privacy are discussed, along with suggestions for future research aimed at improving the accuracy and practicality of fatigue detection systems.

Keywords: Driver Fatigue, Deep Learning, Image-based Fatigue Identification, Physiological Signal-based Fatigue Identification.

1 Introduction

The ubiquitous problem of driver fatigue jeopardizes road safety, contributing to a substantial number of vehicle accidents annually. Early detection of fatigue can prevent many of these accidents, making the development of effective fatigue detection systems a priority in traffic safety research. Fatigue detection technologies have evolved from rudimentary manual assessments to sophisticated systems that employ various data sources, including image capture of the driver's face and physiological signal monitoring. Image-based methods leverage advancements in camera technology and facial recognition algorithms to observe signs of fatigue through changes in facial expressions and eye movements. On the other hand, physiological signal-based methods utilize sensors to record biological data such as heart rate, brain activity, and muscle tension, which provide direct indicators of a driver's alertness level. Each method has its own set of advantages and operational

Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_57

challenges, often influenced by environmental conditions and individual differences among drivers. This introduction discusses the technological foundations of these methods, their application in actual situations, as well as the influence of personal and contextual elements on their efficacy.

2 Image-Based Driver Fatigue Recognition

2.1 Background

Image-based fatigue recognition technology is a crucial path for future research in the area of car safety research in recent years. The main technique employed by image-based fatigue detection system is to capture the driver's facial expressions and head movements using the camera. These biometric data are then analyzed to estimate the driver's level of exhaustion. One benefit of this technology is that it is non-intrusive, allowing continuous monitoring of the driving status without direct contact with the driver's body. The accuracy and efficiency of facial recognition and expression analysis have increased thanks to modern image processing techniques like machine learning and deep learning.

There are a variety of ways to implement this technology, ranging from traditional image processing techniques to sophisticated neural networks, and researchers are always exploring new ways to improve recognition accuracy and real-time performance. Convolutional neural networks (CNNs), for instance, are used to identify common signs of weariness in drivers, such as eye closure frequency and yawn frequency. Furthermore, time-series data analysis using Long Short-Term Memory (LSTM) networks enables the prediction of patterns in driver weariness.

With the development of embedded systems and computing technologies, image-based fatigue monitoring systems can be integrated into modern automobiles to deliver feedback on the status of drivers in real time, thereby enhancing driving safety. These systems show great potential for improving driving safety and are expected to be more widely used globally in the future as the technology is further developed and optimized.

2.2 Representative Works of Image-Based Driver Fatigue Recognition

Xiao et al. presented a sophisticated technique based on a single sample condition that uses facial image recognition to identify driver weariness. This method involves several stages: camera calibration utilizing Zhang Zhengyou's technique, optimal camera parameter determination via linear simulation, and image refinement using the maximum likelihood approach. Subsequently, symmetry algorithms enhance detection efficiency while texture mapping techniques improve image realism. The methodology culminates in facial image recognition, demonstrating a significant improvement in detection accuracy—20% higher than conventional methods—and a reduction in average detection time by over 30% [1].

Bulygin and Kashevnik emphasized the importance of long-term physiological monitoring of drivers, proposing a reference model for image-based fatigue detection aimed at providing timely warnings to drivers to prevent collisions [2]. A methodology for identifying driver weariness was established by Khan et al. using clever facial expression analysis. This framework adeptly handles multi-scale images and addresses challenges such as noise, varying illumination, image scaling, and redundant data, surpassing other similar technologies in classification accuracy [3].

Based on visual signals, Tang et al. developed a real-time online driver tiredness monitoring system. This system automatically monitors fatigue by analyzing the driver's eye and mouth state, employing a detector for facial recognition and Multi-scale Block Local Binary Patterns (MB-LBP) features for precise eye location. In order to track mouth and eye movements and accurately assess the level of weariness, the system also integrates a Kalman filter [4]. A multi-task hierarchical CNN-based fatigue detection system with multi-scale pooling (MSP-Net) was presented by Gu et al. This system has been successfully developed on an embedded platform, integrating face, eye, and mouth state detection. It has proven to have strong performance in challenging driving settings [5].

A method for identifying tiredness driving as an image-based sequence recognition challenge was reported by Xiao et al. It combined long short-term memory with convolutional neural networks. This method employs a deep cascading multitasking framework to extract eye regions from videos, subsequently analyzing the spatial-temporal relationships between frames to predict driver states with enhanced accuracy [6]. An image-based machine learning technique was presented by Singh et al. to lessen user fatigue in interactive model calibration systems. Their innovative approach minimizes the need for user interactions with experts by using unsupervised clustering to group potential solutions based on spatial similarity, further refining the predictive capability of the system [7].

2.3 Summary

Modern image processing techniques—particularly machine learning and deep learning methods—have greatly improved the efficiency and accuracy of facial identification and emotion analysis. For example, the use of CNNs can recognize a driver's blinking frequency and yawning frequency, which are common indicators of fatigue. A combination of LSTM can analyze time-series data to further predict driver fatigue trends. The application of these advanced techniques makes the fatigue detection system not only highly accurate but also responsive and capable of real-time fatigue monitoring.

However, image-based fatigue detection techniques also have some challenges and limitations. For example, different lighting conditions and different facial features of drivers may affect the accuracy of the system. Additionally, processing and analyzing large amounts of data, as well as ensuring the stability and reliability of the system, are also issues that need to be addressed in current research. Nevertheless, image-based fatigue detection technology is still a very promising research direction for future vehicle safety systems.

3 Signal-Based Driver Fatigue Recognition

3.1 Background

Signal-based fatigue recognition is a technique that utilizes physiological signals to monitor and assess an individual's fatigue status. Fatigue manifests itself in a variety of ways, including prolonged reaction time, decreased attention, and weakened judgment. Therefore, the fatigue state can be effectively identified by a variety of physiological and behavioral indicators, such as heart rate, brain waves, eye movements, electromyography, and other signals.

Among them, heart rate variability (HRV) is an important indicator for assessing the autonomic nerve system's activity, which evaluates the body's response to stress by analyzing fluctuations in heart rate. Brainwave signal analysis is another commonly used method for fatigue monitoring, especially analyzing brainwave activity in specific frequency ranges, such as theta and alpha wave activity enhancement, which is usually associated with a fatigue state. Additionally, eye movement parameters, such as blink frequency and changes in pupil diameter, have also been used to assess fatigue.

Fatigue recognition techniques based on physiological signals can monitor an individual's fatigue state in real-time and objectively with higher accuracy and reliability. Therefore, these techniques have been used to develop various fatigue monitoring devices such as wearable devices and in-vehicle systems. These gadgets can keep an eye on the physiological condition of the operator or driver in real time and can alert users when they show signs of weariness, averting potentially dangerous situations. Current research advances in fatigue recognition based on different signals include three main areas: electrocardiographic (ECG), electroencephalographic (EEG), and electromyographic (sEMG) signals.

3.2 Representative Works of Signal-Based Driver Fatigue Recognition

ECG and EEG for Fatigue Identification. ECG and EEG approaches are crucial in the realm of identifying fatigue. By combining feature extraction of ECG signals, Principal Component Analysis (PCA), and machine learning algorithms, Butkevičiūtė et al. successfully developed a new method for everyday mental fatigue recognition, reaching an accuracy of the Random Forest classifier in their experiments exceeding 94.5% [8]. Similarly, Lv et al. explored driver fatigue recognition techniques through EEG signals, demonstrating the effectiveness of deep learning methods in processing complex signals [9]. In addition, Shangguan et al. extracted EEG signal features for driver fatigue state recognition through functional data analysis, with their decision tree classifier achieving an average accuracy of 99.50% on realistically collected driver fatigue EEG signals [10].

Signal Recognition for Muscle Fatigue Identification. The monitoring of muscle fatigue is equally critical and involves different physiological signal processing techniques. Wang et al. utilized sEMG signals with denoising by a modified wavelet thresholding method and the Vslope method to determine the ventilation threshold as an indication of muscle fatigue. With accuracy ranging from 80.33% to 86.69%, the convolutional neural network-support vector machine (CNNSVM) algorithm they created successfully classified muscular tiredness stages [11]. Furthermore, Zeng et al. discovered that A-mode ultrasound (AUS) outperformed sEMG in terms of fatigue robustness when comparing their sensitivity to muscle fatigue in a gesture recognition task [12].

Deep Learning and Multi-Signal Fusion for Fatigue Identification. Multi-signal fusion and deep learning frameworks provide new perspectives for fatigue recognition. Li et al. proposed a hybrid integrated CNN framework based on multiple decomposition methods, which decomposes EEG signals into components through four decomposition methods and learns directly from them, showing an accuracy of 83.48%, demonstrating its potential in various EEG-related tasks [13]. Meanwhile, Karimi and Wang further improved the fatigue recognition performance in noisy environments by constructing fractal dimension and singular value entropy in a functional brain network [14].

3.3 Summary

Signal-based fatigue recognition techniques face some significant challenges that must be addressed to improve their accuracy and reliability. First, inter-individual differences in physiological signals can significantly affect the performance of fatigue recognition systems. Factors such as age, gender, fitness level, and baseline health conditions can influence physiological responses to fatigue, leading to variability in heart rate, brain wave patterns, and muscle activity. These individual differences necessitate the development of personalized models or adaptive algorithms that can account for such variability, ensuring that the system remains accurate across diverse user groups.

Environmental elements that can affect physiological signal stability and quality include temperature, light, and noise. For example, variations in ambient temperature can affect skin conductance and heart rate, while changes in lighting conditions can influence the accuracy of eye-tracking and facial recognition systems. Similarly, background noise can interfere with the recording of audio-based physiological signals, such as speech and respiration rate. These environmental variations present a challenge for maintaining consistent signal quality and require robust preprocessing techniques to filter out noise and artifacts. To address these challenges, future research should focus on several key areas. One important direction is the development of more sophisticated algorithms that can generalize across different individuals and environments. Large, diverse datasets can be used to train machine learning models, especially those that use deep learning techniques, to acquire robust characteristics that are unaffected by changes in the environment and individual differences. Using techniques like domain adaptation and transfer learning can enhance the model's ability to generalize from training data to real-world scenarios.

Another critical area is the integration of multimodal data sources. By combining physiological signals with other types of data, such as behavioral and contextual information, researchers can develop more comprehensive models of fatigue. Multimodal systems can leverage the strengths of each data type, compensating for the weaknesses of any single modality. For instance, while heart rate variability might be influenced by temperature, combining it with facial expression analysis can provide a more reliable indicator of fatigue.

Moreover, adaptive and personalized fatigue recognition systems hold promise for improving accuracy. These systems can learn from individual user data over time, adjusting their parameters to better reflect personal physiological patterns and responses. Such adaptability can be achieved through techniques like online learning and user-specific calibration, which allow the system to fine-tune its performance continually. Finally, rigorous validation and testing in diverse real-world conditions are essential for developing robust signal-based fatigue recognition systems. Field studies and large-scale deployments can provide valuable insights into the system's performance in various settings, highlighting areas for improvement and ensuring that the technology is both practical and effective.

In summary, while signal-based fatigue recognition techniques face significant challenges due to individual differences and environmental factors, future research focused on algorithmic improvements, multimodal data integration, personalized adaptation, and extensive real-world validation can enhance the generalizability and robustness of these systems. By addressing these issues, researchers can develop more accurate and reliable fatigue recognition technologies that contribute to safer and more efficient operations in various applications.

4 Exploiting Multi-Modal Data for Driver Fatigue Recognition

4.1 Background

Multimodal techniques involve the simultaneous use of multiple types of data or information sources to enhance a system's understanding and decision-making capabilities. In the field of artificial intelligence and machine learning, multimodal approaches typically combine data from different modalities such as text, images, audio, and video to gain more comprehensive insights than a single modality. The core of this technique lies in the effective integration of features from different sensors or data sources to optimize the information processing flow through early, late, or hybrid fusion techniques. Multimodal systems can more accurately simulate complex real-world environments, improve user interaction experiences, and provide efficient solutions in multiple application areas such as autonomous driving, sentiment analysis, and health monitoring. This ability to synthesize data across domains is a major advantage of multimodal technologies, allowing them to excel in handling highly complex and varied tasks.

Multimodal fatigue recognition models improve the accuracy and reliability of fatigue state recognition by integrating multiple sensor data. These models combine physiological signals (e.g., electrocardiogram, electroencephalogram, photoplethysmography), behavioral data (including facial expression, eye tracking, and head movement), and environmental data (e.g., vehicle operation data and sound images of the driving environment) to comprehensively reflect the physiological and behavioral states of drivers. Key information is retrieved from numerous raw data sources during the feature extraction step. These features are then successfully integrated using feature fusion approaches (e.g., early fusion, late fusion, or hybrid fusion). Lastly, the model is trained using machine learning or deep learning techniques, such as support vector machines, random forests, and deep neural networks, to identify tiredness states from composite data. This multimodal approach not only enhances the adaptability of the model but also improves the efficiency and accuracy of recognizing fatigue in complex and dynamic environments.

4.2 Representative Works of Multi-Modal Driver Fatigue Recognition

Non-Invasive Sensors and Feature Selection for Driving Fatigue Recognition. Henni et al. improved the accuracy of driving fatigue detection by using non-invasive sensors, analyzing features associated with driving behavior and facial expression changes, and feature selection methods and meta-analysis. When Support Vector Machines and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) classifiers were used to real data of 66 senior drivers, the selected attributes improved the classification accuracy to 89.13%. [15]. Furthermore, Kong et al. presented a non-invasive multimodal fusion fatigue detection method that is superior to conventional techniques, achieving 98.2% accuracy on the Microbe-Drug Association Database (MDAD) dataset. Motion tracking convolutional neural networks with bi-directional long- and short-term memory networks and remote photoplethysmography form the basis of this technique [16].

Multimodal Fusion Techniques for Fatigue and Attention Recognition. The application of multimodal fusion techniques is opening up new avenues for fatigue and attention recognition. A multimodal Bayesian network-based driving fatigue and distraction assessment system assembles multiple sensors in a driving simulator, processes the features independently through visual, audio, and signal modules, and then fuses these features to estimate driver attention. The system's accuracy for fatigue detection is 98.4%, and its accuracy for distraction detection is 90.5%, according to the trial data [17]. Furthermore, Zhou et al. showed the effectiveness of multimodal feature fusion in efficiently monitoring fatigue by the proposal of a Bayes-gcForest-based multimodal feature fusion brain fatigue recognition system employing EEG and ECG data [18].

Future-Proof Multimodal and Federated Learning Approaches for Driving Fatigue Detection. Yu et al. suggested a federated annotation technique based on multimodal sensory data for driving fatigue detection in response to the growth of multimodal sensory data. This method effectively utilizes the correlation and complementary information between multiple modalities and shows superior performance to other methods even with only 10% of labeled data [19]. Additionally, Ortega et al. presented a large-scale multimodal driving monitoring dataset, containing real and simulated driving scenarios for the development of a driving behavior recognition system, demonstrating its high accuracy and efficiency on a cost-effective Central Processing Unit (CPU) platform [20].

4.3 Summary

The application of multimodal models in fatigue recognition significantly enhances the precision and dependability of fatigue state recognition by integrating multiple sensors and data sources. These models leverage data from various modalities, such as visual cues, physiological signals, and environmental factors, to provide a comprehensive assessment of a driver's state. By combining information from facial expressions, eye movements, heart rate variability, and other biometric signals, multimodal systems can detect subtle indicators of fatigue that single-modality systems might miss. These models are highly adaptable to complex environments, allowing them to function effectively in a variety of real-world scenarios. For instance, in autonomous driving, multimodal fatigue recognition systems can seamlessly integrate with vehicle control systems to provide real-time alerts and interventions, enhancing overall safety. In sentiment analysis, these models can be employed to monitor user engagement and emotional state, providing valuable insights for improving user interaction experiences. The potential applications of multimodal fatigue recognition extend beyond these fields. In healthcare, such systems can monitor patients' alertness and well-being, particularly in high-stress environments like intensive care units. In occupational safety, they can be used to ensure that workers operating heavy machinery or performing critical tasks remain alert and focused, thereby reducing the risk of accidents.

However, the implementation of multimodal models is relatively complex in terms of data integration and processing. The need to synchronize and analyze data from diverse sources requires sophisticated algorithms and robust computational infrastructure. These systems need to be able to process massive amounts of data quickly, which can be resource-intensive. Additionally, the variability in data quality and sensor performance across different conditions poses a challenge for maintaining consistent accuracy. Model training and optimization in multimodal systems also present significant difficulties. Ensuring that the models generalize well across different populations and environments requires extensive and diverse training datasets. The integration of multiple data types necessitates advanced techniques in feature extraction, data fusion, and machine learning. Researchers must also address potential overfitting issues, ensuring that models perform well not only on training data but also in real-world applications.

Moreover, the use of sensitive physiological and behavioral data introduces critical privacy and security concerns. Protecting user data and ensuring compliance with regulations such as General Data Protection Regulation (GDPR) are paramount. Researchers and developers must implement robust encryption and anonymization techniques to safeguard data and maintain user trust. Despite these challenges, multimodal models, with their cross-domain integration capabilities, show strong potential in dealing with dynamic and variable tasks. They represent a significant advancement in fatigue recognition technology, promising to enhance safety and efficiency across multiple domains. Future studies should concentrate on making these systems more resilient and scalable, creating more effective data processing algorithms, and resolving moral issues with data privacy. By overcoming these hurdles, multimodal fatigue recognition systems can achieve widespread adoption and make a meaningful impact on public safety and well-being.

5 Conclusion

Detecting driver fatigue through technological means is an increasingly vital component of road safety strategies. While both image-based and physiological signal-based methods provide valuable insights into a driver's state, they also face significant challenges that can hinder their effectiveness. Image-based systems must contend with variations in lighting and the driver's facial orientation, while physiological methods require the management of sensor placement and the potential discomfort it may cause. Moreover, issues like data privacy and the need for robust, real-time processing capabilities remain critical concerns. Subsequent investigations ought to concentrate on augmenting the precision and dependability of these systems via enhanced algorithms, superior sensor technologies, and more advanced data integration. Ultimately, the goal is to develop unobtrusive, accurate, and reliable systems that can be widely implemented to significantly reduce fatigue-related accidents on roads.

References

1. Xiao, C., Han, L., & Chen, S.: Automobile Driver Fatigue Detection Method Based on Facial Image Recognition under Single Sample Condition. Symmetry, 13(7), 1195 (2021).
2. Bulygin, A., & Kashevnik, A.: Image-Based Fatigue Detection of Vehicle Driver: State-of-the-Art and Reference Model. In 30th Conference of Open Innovations Association FRUCT. pp. 24-31. IEEE, Oulu (2021).
3. Khan, S. A., Hussain, S., Xiaoming, S., & Yang, S.: An effective framework for driver fatigue recognition based on intelligent facial expressions analysis. IEEE Access, 6, 67459-67468 (2018).
4. Tang, X., Zhou, P., & Wang, P.: Real-time image-based driver fatigue detection and monitoring system for monitoring driver vigilance. In 35th Chinese control conference. pp. 4188-4193. IEEE, Yichang (2016).
5. Gu, W. H., Zhu, Y., Chen, X. D., He, L. F., & Zheng, B. B.: Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model. IET Image Processing, 12(12), 2319-2329 (2018).
6. Xiao, Z., Hu, Z., Geng, L., Zhang, F., Wu, J., & Li, Y.: Fatigue driving recognition network: fatigue driving recognition via convolutional neural network and long short-term memory units. IET Intelligent Transport Systems, 13(9), 1410-1416 (2019).

7. Singh, A., Minsker, B. S., & Bajcsy, P.: Image-based machine learning for reduction of user fatigue in an interactive model calibration system. Journal of computing in civil engineering, 24(3), 241-251 (2010).
8. Butkevičiūtė, E., Michalkovič, A., & Bikulčienė, L.: Ecg signal features classification for the mental fatigue recognition. Mathematics, 10(18), 3395 (2022).
9. Lv, C., Nian, J., Xu, Y., & Song, B.: Compact vehicle driver fatigue recognition technology based on EEG signal. IEEE transactions on intelligent transportation systems, 23(10), 19753-19759 (2021).
10. Shangguan, P., Qiu, T., Liu, T., Zou, S., Liu, Z., & Zhang, S.: Feature extraction of EEG signals based on functional data analysis and its application to recognition of driver fatigue state. Physiological Measurement, 41(12), 125004 (2020).
11. Wang, J., Sun, Y., & Sun, S.: Recognition of muscle fatigue status based on improved wavelet threshold and CNN-SVM. IEEE Access, 8, 207914-207922 (2020).
12. Zeng, J., Zhou, Y., Yang, Y., Yan, J., & Liu, H.: Fatigue-sensitivity comparison of sEMG and A-mode ultrasound based hand gesture recognition. IEEE Journal of Biomedical and Health Informatics, 26(4), 1718-1725 (2021).
13. Li, R., Gao, R., Yuan, L., Suganthan, P. N., Wang, L., & Sourina, O. An enhanced ensemble deep random vector functional link network for driver fatigue recognition. Engineering Applications of Artificial Intelligence, 123, 106237 (2023).
14. Hamid Reza Karimi and N. Wang, Sensor Systems and Software. Springer Nature, (2023).
15. Deng, P. Y., Qiu, X. Y., Tang, Z., Zhang, W. M., Zhu, L. M., Ren, H., et al.: Detecting fatigue status of pilots based on deep learning network using EEG signals. IEEE Transactions on Cognitive and Developmental Systems, 13(3), 575-585 (2020).
16. Kong, L., Xie, K., Niu, K., He, J., & Zhang, W.: Remote photoplethysmography and motion tracking convolutional neural network with bidirectional long short-term memory: Non-invasive fatigue detection method based on multi-modal fusion. Sensors, 24(2), 455 (2024).
17. Multi-Modal Driver Distraction and Fatigue Detection and Warning System, Phase II (SBI R) | FMCSA, Dot.gov, URL: https://www.fmcsa.dot.gov/research-and-analysis/research/m ulti-modal-driver-distraction-and-fatigue-detection-and-warning. Last accessed May 19, 20 24.
18. Zhou, Y., Chen, P., Fan, Y., & Wu, Y.: A Multimodal Feature Fusion Brain Fatigue Recognition System Based on Bayes-gcForest. Sensors, 24(9), 2910 (2024).
19. Yu, S., Yang, Q., Wang, J., & Wu, C.: FedUSL: A Federated Annotation Method for Driving Fatigue Detection based on Multimodal Sensing Data. ACM Transactions on Sensor Networks, 1550-4859 (2024).
20. Ortega, J. D., Kose, N., Cañas, P., Chao, M. A., Unnervik, A., Nieto, M., ... & Salgado, L.: Dmd: A large-scale multi-modal driver monitoring dataset for attention and alertness analysis. In The European Conference on Computer Vision Workshops. pp. 387-405. Springer, Glasgow (2020).