Journal of Electronic Research and Application, 2025, Volume 9, Issue 4
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Detecting Drowsiness Behind the Wheel: A Lightweight Approach Based on Eye and Mouth Aspect Ratios

Heyang Ni*

Department of Physics, University of California, Santa Barbara 93106, United States

**Author to whom correspondence should be addressed.*

Abstract: Driver distraction is a leading cause of traffic accidents, with fatigue being a significant contributor. This paper introduces a novel method for detecting driver distraction by analyzing facial features using machine deep learning and 68 face model. The proposed system assesses driver tiredness by measuring the distance between key facial landmarks, such as the distance between the eyes and the angle of the mouth, to evaluate signs of drowsiness or disengagement. Real-time video feed analysis allows for continuous monitoring of the driver's face, enabling the system to detect behavioral cues associated with distraction, such as eye closures or changes in facial expressions. The effectiveness of this method is demonstrated through a series of experiments on a dataset of driver videos, which proves that the approach can accurately assess tiredness and distraction levels under various driving conditions. By focusing on facial landmarks, the system is computationally efficient and capable of operating in real-time, making it a practical solution for in-vehicle safety systems. This paper discusses the system's performance, limitations, and potential for future enhancements, including integration with other in-vehicle technologies to provide comprehensive driver monitoring.

Keywords: Driver drowsiness detection; Eye aspect ratio (EAR); Mouth aspect ratio (MAR); Facial landmark detection; Real-time monitoring

Online publication: August 7, 2025

1. Introduction

Among all of the facial expressions that can be detected from human faces, distraction is one expression that is being researched and explored the most, not only because the appearance of distraction may lead to severe outcomes, but also because the blurred definition of distraction makes the computer hard to decide whether humans are truly distracted. Generally, the basic definition of distraction can be separated into four categories: vision distraction, manual distraction, cognitive distraction, and auditory distraction [1]. In this paper, vision distraction and manual distraction will be discussed as video detection of driving distraction is the main focus of the research being conducted. In the context of driving, distraction recognition is usually connected with fatigue; therefore, understanding driving fatigue has become more crucial in recent years due to the fast development of AI technology. By using a camera for capturing images of different parts of the face such as eyes and mouth, and by integrating machine learning and computer vision technologies, the software is able to detect and analyze the emotional states of drivers and take certain actions based on the severity of fatigue to keep drivers concentrated. Fatigue driving, also known as drowsy driving, refers to operating a motor vehicle while experiencing physical or mental exhaustion due to inadequate sleep, prolonged driving, or other physiological and psychological stressors [2]. This condition significantly impairs reaction time, attention, and decision-making ability, thereby increasing the risk of road accidents [2]. This condition leads to decreased attention, slower reaction times, impaired decision-making, and an increased likelihood of accidents. Fatigue driving is particularly dangerous because, unlike driving under the influence of alcohol or drugs, it often goes unnoticed by the driver until it significantly affects their ability to control the vehicle. According to a US survey, about 20 percent of vehicle crashes are due to driving fatigue [3]; therefore, understanding what causes driving fatigue is significantly important not only for driver safety but also for road safety.

In order to utilize and integrate that knowledge into software, scientists have made efforts for decades. In recent years, machine deep learning has become a more commonly used tool for fatigue expression recognition. Machine deep learning has revolutionized driving fatigue detection by enabling systems to automatically identify and verify individuals based on facial features. Specifically, it is widely used in this field due to its ability to automatically learn hierarchical features from raw image data, such as edges, textures, and patterns, which are crucial for recognizing faces. These models are trained on large datasets of facial images, enabling them to generalize well and handle variations in lighting, pose, and occlusions. Deep neural networks can also leverage transfer learning, where pre-trained models on vast datasets are fine-tuned for specific face recognition tasks, improving accuracy and efficiency. The advancements in deep learning have significantly improved the robustness and speed of fatigue recognition systems. However, deep machine learning also faces problems such as GPU inefficiency, lack of annotated data sets, and heavy work in collecting and categorizing the data.

In this paper, the author introduces a new way of fatigue recognition established on a model that has already been trained by supervising human eyes and mouth. My approach involves developing code that further trains the computer on fatigue detection based on eyes and mouth conditions in way of deep learning, and adding more data to increase the precision of the system. This dataset includes labeled images of drivers in various states of alertness, enabling the model to learn patterns associated with drowsiness. The essay explores the development of facial expression detection, the transition from traditional to modern machine learning techniques, and how my new approach builds upon these advancements.

2. Literature review

Fatigue detection has evolved significantly over the past few decades, transitioning from simple hand-crafted feature-based techniques to sophisticated deep learning-based approaches. Deep learning has revolutionized facial expression detection, enabling automated feature extraction and improving classification accuracy. The introduction of large-scale annotated datasets, such as AffectNet and FER2013, further enhanced the performance of deep learning models [4]. In recent years, driving fatigue detection has been researched intensively, and researchers have proposed a bunch of methods and measures to determine the level of fatigue to prevent traffic accidents. Recent research has focused on developing and enhancing fatigue detection techniques using various approaches, including image-based monitoring, physiological signal analysis, vehicle-based detection, and hybrid systems.

Image-based detection methods utilize cameras and computer vision to monitor drivers' facial features, such as eye movements, yawning, and head positions, to assess drowsiness levels. Deep learning techniques have improved the accuracy and robustness of these systems under different lighting conditions. A study by Li developed a driver fatigue detection system based on deep learning models that analyze facial landmarks and detect fatigue indicators, such as prolonged eye closure and yawning [5]. There are also recent studies that leveraged architectures such as ResNet and Vision Transformers (ViTs) to achieve state-of-the-art results in facial expression recognition [6]. These models benefit from large-scale pretraining on diverse datasets, improving generalization across different demographics and lighting conditions. Integrating emotion recognition with facial analysis has also been shown to improve detection accuracy, as a recent study combined convolutional neural networks (CNNs) with emotional state analysis to enhance system robustness.

More recently, deep learning models such as Convolutional Neural Networks (CNNs) and hybrid models combining CNNs with Recurrent Neural Networks (RNNs) like Long Short-Term Memory (LSTM) networks have been introduced to enhance robustness against variability in lighting, facial structure, and occlusions [7]. These models can automatically extract spatial and temporal features from video frames, allowing more sophisticated detection of fatigue patterns without relying on handcrafted features. However, while deep learning methods significantly improve accuracy and adaptability, they come with substantial limitations.

First, deep learning models require large, diverse, and well-labeled datasets for effective training. Annotating fatigue-related behaviors is labor-intensive, particularly in real-world driving conditions where ethical and safety concerns limit data collection [8]. Additionally, such models are computationally intensive, often necessitating GPU acceleration for real-time deployment, which may not be feasible for embedded or mobile systems. Another significant challenge is generalization: deep models trained on specific populations or controlled environments often struggle with real-world variability such as changes in camera angle, illumination, or occlusions like sunglasses and facial hair.

Moreover, deep learning systems are often black boxes, providing little insight into their decision-making process, which poses challenges for trust, interpretability, and debugging—particularly in safety-critical applications like autonomous driving. There is also an increasing concern over algorithmic bias, as models trained on demographically limited datasets may underperform on underrepresented groups, leading to disparities in fatigue detection accuracy [9].

In summary, behavior-based fatigue detection using deep learning offers promising advancements but is accompanied by significant limitations related to data requirements, computational costs, generalizability, and fairness. Addressing these challenges is crucial for deploying reliable and ethical fatigue detection systems in practical settings.

3. Methodology

This study presents a hybrid approach to detecting driving fatigue, combining computer vision techniques with supervised machine learning. The workflow consists of three key stages: image data collection, feature extraction using 68 face model, and classification using a decision tree model. The main approach to detecting fatigue when driving is through deciding the opening and closing of the eyes and the opening degree of the mouth.

3.1. Image data collection

Images of drivers were collected under various driving conditions using an in-cabin camera system. The dataset captures real-time facial expressions, specifically focusing on the eyes and mouth, as these are critical indicators of fatigue (e.g., eye closure, yawning). The images were annotated with binary labels—fatigued or alert—based on driver behavior observed during the data collection process or taken from publicly available drowsiness datasets. Moreover, the dataset also includes videos of drivers under various driving conditions for unexpected behaviors' influence on precision.

To ensure diversity, the dataset includes drivers with different facial features, lighting conditions, and camera angles. Images were preprocessed by resizing, grayscale conversion, and histogram equalization to enhance contrast and ensure uniformity.

3.2. Feature extraction using 68 face model

Facial features were identified using the 68-point face landmark model implemented via the Dlib library. This model maps 68 specific points on the human face, and based on an ensemble of regression trees, which allows for real-time detection of facial structures without requiring deep learning or GPU acceleration. It outputs the (x, y) coordinates of facial landmarks for each detected face.

From the 68 facial landmarks, two primary features were extracted for fatigue assessment:

(1) Eye Aspect Ratio (EAR): Calculated using the distances between selected eye landmarks to determine the degree of eye openness. A consistently low EAR indicates prolonged eye closure—a strong indicator of drowsiness.

(2) Mouth Aspect Ratio (MAR): Computed from mouth landmarks to detect yawning. Frequent or prolonged mouth opening suggests driver fatigue. These features were calculated frame-by-frame.

3.3. Fatigue detection method based on EAR and MAR

The proposed fatigue detection system utilizes two geometric indicators extracted from facial landmarks: the Eye Aspect Ratio (EAR) and the Mouth Aspect Ratio (MAR). EAR is employed to quantify eye openness, while MAR measures the extent of mouth opening. These ratios are derived from a 68-point facial landmark model, which identifies precise facial features such as eyelids and lips.

The EAR is calculated based on the distances between specific vertical and horizontal eye landmarks as follows:

$$EAR = \frac{\frac{\|P_{leftup} - P_{leftdown}\|}{P_{lefthorizontal}} + \frac{\|P_{rightup} - P_{rightdown}\|}{P_{righthorizontal}}}{2} \tag{1}$$

When the eyes are open, vertical distances are relatively large, resulting in a high EAR. As the eyes close, vertical distances decrease and EAR drops accordingly. However, due to natural differences in eye shapes among individuals, using a fixed EAR threshold can lead to misclassification, especially for users with naturally small eyes. To address this issue, a dynamic EAR threshold is introduced. During an initial calibration phase, the system tracks the minimum and maximum EAR values over time. The personalized threshold is computed using:

$$ThresholdEAR = EARmin + \alpha \cdot (EARmax - EARmin) \tag{2}$$

where α is a tunable hyperparameter (e.g., 0.25). This threshold is used to determine eye closure in real-time frames. After determining the closure of eyes, fatigue level is determined by evaluating the duration of consecutive frames in which the EAR remains below a dynamic closure threshold. Two distinct fatigue levels are defined:

(1) Light fatigue: Continuous eye closure lasting at least 1.5 seconds.

(2) Severe fatigue: Continuous eye closure lasting at least 2.5 seconds.

These thresholds are supported by prior research. Wierwille and Ellsworth identified eye closure durations exceeding 1.5 seconds as indicative of early drowsiness and cognitive disengagement, beyond the typical blink duration of approximately 300–400 milliseconds [10]. Further, Dinges established that eye closures exceeding 2.0 seconds correlate strongly with microsleep episodes and degraded driving performance, justifying the use of a 2.5-second threshold for severe fatigue [11].

In the implementation, binary eye state values (1 = closed, 0 = open) are recorded in a rolling queue at a frame rate of 30 frames per second. A sliding window of 90 frames (3 seconds) is used to monitor these states. Light fatigue is flagged when 45 or more consecutive frames indicate eye closure (1.5 seconds), while severe fatigue is identified after 75 or more consecutive closed-eye frames (2.5 seconds). This temporal classification framework enables non-intrusive, camera-based assessment of fatigue with high temporal precision [12,13].

Similarly, the MAR is calculated using vertical and horizontal distances between upper and lower lip landmarks, additionally, calculate the MAR for both inner lip points and outer lip points:

$$MAR = \frac{1}{2} \cdot \left(\frac{\|P_{uplip} - P_{downlip}\|}{3 \cdot P_{insidehorizontal}} + \frac{\|P_{uplip} - P_{downlip}\|}{3 \cdot P_{outhorizontal}}\right) \quad (3)$$

To detect yawning, a dynamic MAR threshold is computed in the same way:

$$\text{ThresholdMAR} = \text{MARmin} + \beta \cdot (\text{MARmax} - \text{MARmin}) \quad (4)$$

where β is another tunable parameter (e.g., 0.5). A yawning event is identified when the MAR remains above this threshold for a sustained number of frames, typically corresponding to 4–6 seconds (e.g., 60 out of 90 frames at 30 FPS), which reflects the natural duration of a yawn.

To evaluate fatigue, two binary queues—queue_eye and queue_mouth—are maintained to record the recent states of eye closure and mouth opening, respectively. These queues have a fixed length representing a sliding window (e.g., 3 seconds). Fatigue level is determined based on the number of frames within these windows that indicate eye closure or yawning. Specifically, the system classifies the fatigue status as follows:

(1) Level 0 (Normal): No significant signs of fatigue are detected.

(2) Level 1 (Mild fatigue): Eye closure exceeds a mild threshold (e.g., 30% of the window) or yawning is detected.

(3) Level 2 (Severe fatigue): Prolonged eye closure (e.g., over 60% of the window) or simultaneous eye closure and yawning is detected.

This approach enables real-time, frame-level detection of fatigue symptoms without requiring additional sensors, and adapts effectively to diverse facial structures and behaviors. The entire system is depicted in **Figure 1**. Some examples of image detection are provided in **Figure 2**.

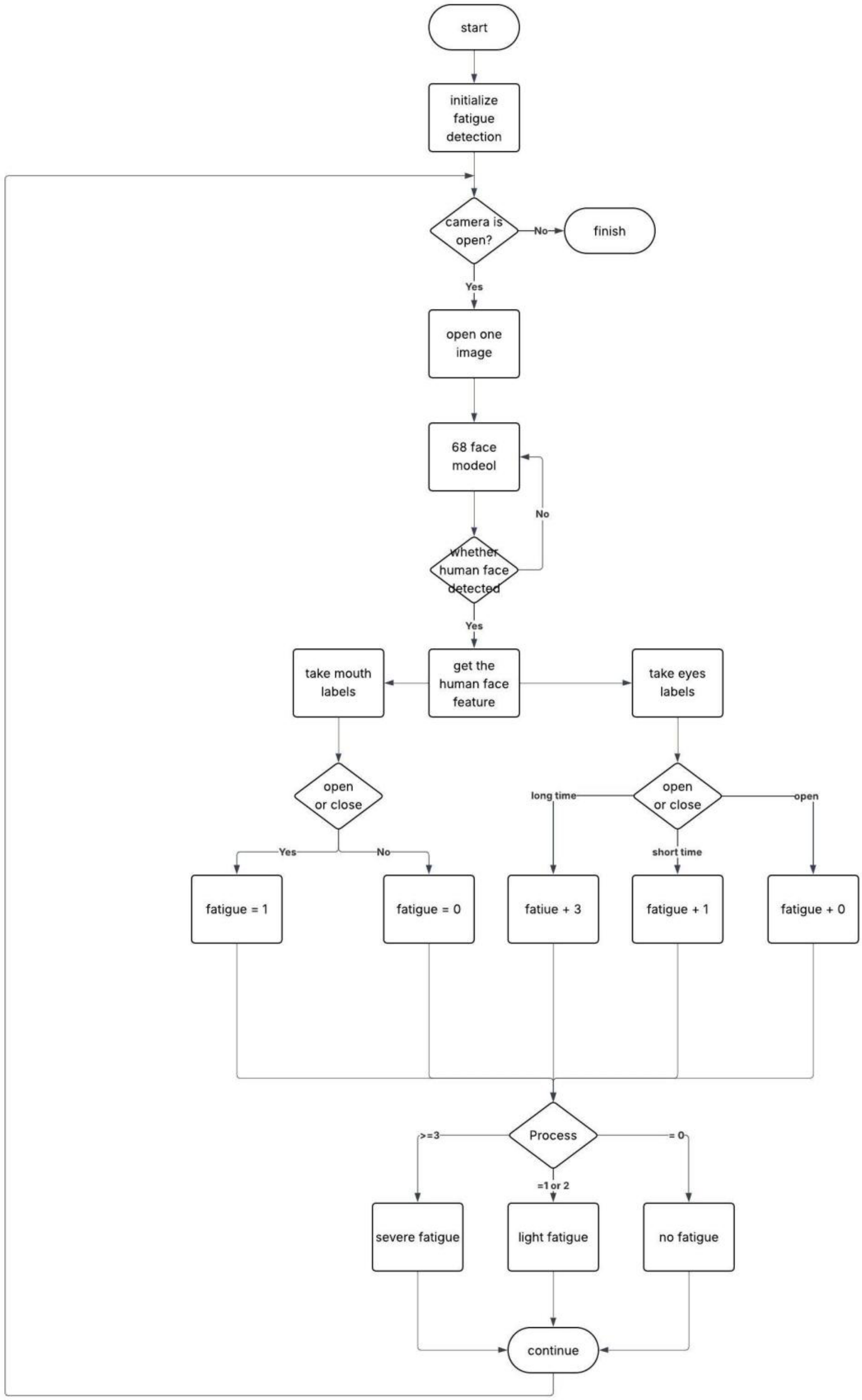
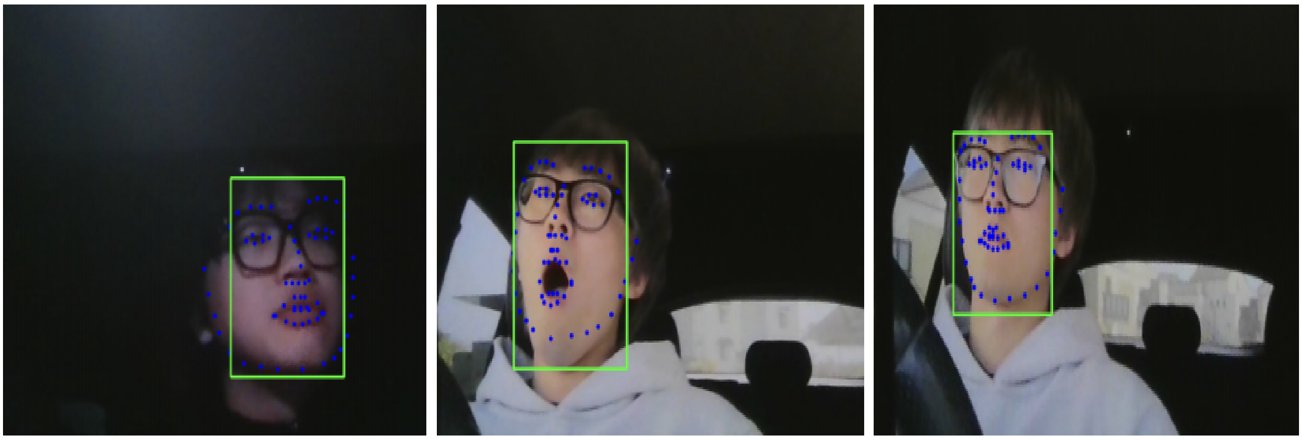


Figure 1. Fatigue detection flow chart

4. Result analysis

The proposed fatigue detection system was evaluated under different lighting conditions and user characteristics to assess its robustness and accuracy (**Table 1**). In daytime scenarios with sufficient ambient lighting, the system achieved a high accuracy rate of 99% in detecting both eye closure and yawning events from 46 volunteers with an average of 2 minutes of video. However, in low-light or nighttime conditions, the accuracy dropped to 95% at the same sample size, likely due to insufficient illumination affecting the precision of facial landmark detection. This performance gap suggests that integrating infrared or night-vision cameras could significantly enhance detection reliability in dark environments.



(a) (b) (c)

Figure 2. (a) no fatigue; (b) light fatigue; (c) severe fatigue

Furthermore, during testing with drivers who naturally have smaller eyes, the system's fatigue detection accuracy decreased substantially. This decline is attributed to the limitations of the camera resolution, which affects the precision of eye aspect ratio (EAR) calculation when vertical eyelid distances are already minimal. Improving camera resolution or employing higher-fidelity sensors may help resolve this issue.

Table 1. Accuracy of fatigue detection under different conditions

Testing condition	Accuracy	Possible limitation	Possible improvement
Daytime (bright lighting)	99%	None	—
Nighttime (low lighting)	95%	Poor illumination affects landmark precision	Use infrared or night-vision camera
Small-eyed driver (daytime)	83%	Eye landmarks less distinguishable at low resolution	Use higher-resolution cam- era or zoom lens
Fast blinking (controlled)	98%	May cause false fatigue trigger in edge cases	Temporal smoothing or blink classification
Mouth-only yawning (no eye closure)	97%	Yawn may be brief or partially occluded	Use mouth duration tracking + MAR sequence

The current system was tested on a machine equipped with an NVIDIA GeForce RTX 3070 GPU and 16GB of RAM. While the webcam delivers frames at 30 frames per second (FPS), the effective processing throughput of the fatigue detection pipeline is approximately 10 FPS, constrained by the computational demands of face detection, landmark extraction, and fatigue classification. With increased computational power or optimized lightweight models, the system's real-time performance and accuracy could be further improved, making it more suitable for deployment in resource-constrained embedded systems or edge devices.

5. Discussion and conclusion

The results of this study demonstrate that combining deep learning-based feature extraction with interpretable machine learning models can effectively detect signs of driver fatigue. By focusing on visual indicators—specifically eye closure and mouth openness—extracted from real-time facial images, the system is capable of identifying early signs of drowsiness such as frequent blinking, prolonged eye closure, and yawning. These visual cues are widely recognized as reliable indicators of fatigue, and their effectiveness has been validated in prior studies [7,14].

Unlike approaches that rely on deep learning methods such as convolutional neural networks (CNNs), the 68-point model used in this study is based on a lightweight regression-tree-based method. This allows for fast, resource-efficient facial analysis without requiring large training datasets or significant computational power. The use of the facial landmark model ensures robust performance under normal lighting and facial visibility conditions, making it suitable for real-time deployment in driver assistance systems.

The extracted features were input into a decision tree classifier, which provided clear and interpretable decision paths for distinguishing between fatigued and alert states. The decision tree model showed strong performance in terms of accuracy and recall, indicating that the selected facial features are highly predictive of fatigue-related behaviors. Moreover, the transparency of the decision tree allowed for insight into which features—such as eyelid closure ratio and mouth aspect ratio—had the greatest influence on classification outcomes.

However, an important limitation of the current approach is its reliance on facial visibility. In scenarios where the driver's face is partially obstructed (e.g., by sunglasses or occlusion), feature extraction becomes unreliable. Previous studies have addressed this issue by incorporating additional physiological or vehicle-based data, such as heart rate variability or steering wheel behavior [15]. Integrating such multi-modal data sources could enhance the robustness and accuracy of the detection system. Furthermore, while the decision tree model is computationally efficient and interpretable, it may lack the predictive power of more complex models such as random forests or gradient boosting machines, which could be explored in future work. Moreover, while this study used a labeled dataset, real-world deployment would require the system to adapt to drivers with different baseline behaviors. Implementing a personalized calibration phase, or leveraging unsupervised learning techniques, may help mitigate this variability.

In conclusion, this study highlights the effectiveness of combining deep learning with interpretable models for fatigue detection. The proposed fatigue detection system holds significant potential for deployment in real-world transportation scenarios, including various types of professional driving, such as public buses, taxis, and long-haul freight trucks, where driver fatigue is a major contributor to road accidents. In public transportation, the system can continuously monitor the driver's alertness in real time and issue early warnings when signs of light or severe fatigue are detected, helping prevent potential collisions or operational errors. For ride-hailing and taxi fleets, fatigue monitoring can be incorporated into dashboard cameras or mobile applications to ensure passenger safety and enforce safe driving behavior. Future work may focus on improving low-light performance, optimizing computational efficiency, and expanding multi-modal integration to further increase reliability and adaptability in diverse driving environments.

Disclosure statement

The author declares no conflict of interest.

References

[1] Abtahi S, Omidyeganeh M, Shirmohammadi S, et al., 2014, Yawning Detection Using Embedded Smart Cameras. IEEE Transactions on Instrumentation and Measurement, 63(7): 1842–1854.

[2] Teletrac Navman, n.d., What Is Driver Fatigue? viewed April 13, 2025, https://www.teletracnavman.com/fleet-management-software/safety/resources/what-is-driver-fatigue

[3] Sprajcer M, Dawson D, Kosmadopoulos A, et al., 2023, How Tired Is Too Tired to Drive? A Systematic Review Assessing the Use of Prior Sleep Duration to Detect Driving Impairment. Nature and Science of Sleep, 15: 175–206.

[4] Mollahosseini A, Hasani B, Mahoor MH., 2017, AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing, 10(1): 18–31.

[5] Li D, Zhang X, Liu X, et al., 2023, Driver Fatigue Detection Based on Comprehensive Facial Features and Gated Recurrent Unit. J Real-Time Image Proc, 20(19).

[6] Kollias D, Schulc A, Hajiyev E, et al., 2021, Deep Affect Prediction In-the-Wild: Aff-Wild Database and Challenge. IEEE Transactions on Affective Computing.

[7] Abtahi S, Omidyeganeh M, Shirmohammadi S, et al., 2014, Yawning Detection Using Embedded Smart Cameras. IEEE Transactions on Instrumentation and Measurement, 63(7): 1842–1854.

[8] Dwivedi R, Biswas D, Chakraborty S, 2014, Real-Time Driver Drowsiness Detection Using Computer Vision. International Journal of Computer Applications, 103(10): 36–39.

[9] Buolamwini J, Gebru T, 2018, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research, 81: 1–15.

[10] Wierwille WW, Ellsworth LA, 1994, Research on Vehicle-Based Driver Status/Performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness. National Highway Traffic Safety Administration.

[11] Dinges DF, Mallis MM, Maislin G, et al., 1998, PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance. U.S. Department of Transportation, Federal Highway Administration.

[12] Zhao Y, Zheng Y, Zhang Y, 2021, Wearable EEG-Based Drowsiness Detection Using Convolutional Neural Networks. Sensors, 21(3): 926.

[13] Sikander G, Anwar S, 2019, Driver Fatigue Detection Systems: A Review. IEEE Transactions on Intelligent Transportation Systems, 20(6): 2339–2352.

[14] Albadawi Y, Takruri M, Awad M, 2022, A Review of Recent Developments in Driver Drowsiness Detection Systems. Sensors, 22(5): 2069.

[15] Corneanu CA, Simn MO, Cohn JF, et al., 2016. Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(8): 1548–1568.