Advances in Computer, Signals and Systems (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2025.090307
ISSN 2371-8838 Vol. 9 Num. 3

Application of YOLO-Based Face Recognition in Fatigue Driving Detection

Jianlei Li[1,a,*], Runyi Hu[1,b]

[1]*School of Mathematics and Statistics, North China University of Water Resources and Electric Power, Zhengzhou, Henan, China*

[a]*lijianlei@ncwu.edu.cn*, [b]*z20231080790@stu.ncwu.edu.cn*

**Corresponding author*

Keywords: Fatigue Driving Detection; YOLOv8; FaceMesh; Face Recognition; Eye Aspect Ratio (EAR); Mouth Aspect Ratio (MAR); Deep Learning

Abstract: Fatigue driving is a major contributing factor to traffic accidents. Accurately and real-time identification of driver fatigue has become a research priority in the field of intelligent driving safety. This paper proposes a face recognition method that integrates YOLOv8 and FaceMesh to achieve high-precision fatigue driving detection. This method first uses the YOLOv8 model to rapidly locate the driver's face. Furthermore, the FaceMesh model is introduced to extract facial key points. Fatigue behavior features such as the eye aspect ratio (EAR) and mouth opening/closing ratio (MAR) are calculated, and state discrimination is performed using time-series statistical logic. Experimental results show that this method achieves 93.4% accuracy, 91.6% recall, and 92.5% F1-score on a public dataset, outperforming the traditional YOLOv5 and keypoint method combination. It also maintains robustness in complex scenarios such as nighttime and occlusion. These results demonstrate the effectiveness and practicality of this method in fatigue driving detection, providing a viable technical path for intelligent vehicle monitoring systems.

1. Introduction

With the continuous development of intelligent transportation and in-vehicle safety systems, driver fatigue has become a major factor in traffic accidents. According to statistics from the World Health Organization and transportation departments of various countries, driver fatigue is a significant contributing factor to highway accidents, and its harmful effects cannot be ignored [1]. Traditional fatigue detection methods rely on physiological sensors (such as EEG, ECG, and EMG) or driving behavior parameters (such as steering wheel deviation and braking frequency) [2]. While these methods offer certain advantages in accuracy, practical deployments face challenges such as inconvenience, low user acceptance, and poor real-time performance. Therefore, achieving non-contact, real-time fatigue detection using computer vision technology has become a hot research topic [3].

In recent years, deep learning-based visual detection technology has made significant progress in driving behavior analysis. Facial recognition, as an important means of obtaining driver status information, plays a key role in fatigue behavior detection. By dynamically monitoring the driver's eyes and mouth, it is possible to automatically identify fatigue characteristics such as blinking, eye closure, and yawning [4]. The widespread application of object detection algorithms in this field provides strong support for real-time detection. The YOLO (You Only Look Once) family of algorithms, in particular, is widely used in tasks such as face and expression detection due to its high speed and high accuracy. YOLOv8, the latest version of this family, boasts significant improvements in both accuracy and real-time performance, providing a superior foundation for fatigue detection [5].

Although YOLOv8 performs well in facial region detection, it still suffers from insufficient resolution and less precise feature localization when detecting smaller local areas, such as the eyes and mouth. To address this challenge, this paper introduces the FaceMesh model for auxiliary feature extraction. FaceMesh regresses 468 3D facial landmarks to achieve high-precision localization of key areas such as the eyes and mouth. Fusion of yolov8 and facemesh allows yolov8 to quickly locate the facial ROI, reducing the FaceMesh computational scope and improving overall efficiency. Furthermore, FaceMesh refines YOLO detection results, thereby improving the accuracy of local fatigue feature detection. This fusion strategy balances detection speed with enhanced local feature extraction, demonstrating promising practical application prospects [6].

In summary, this paper aims to design and implement a facial fatigue feature detection method based on the fusion of YOLOv8 and FaceMesh for use in fatigue driving status recognition. By extracting and analyzing key metrics such as eye closure (EAR) and mouth opening/closing ratio (MAR), combined with a temporal behavior judgment mechanism, this method achieves high-precision fatigue status recognition. This paper will validate the effectiveness of this method using public datasets and real-world acquisition scenarios, analyzing its robustness under varying lighting and occlusion conditions, further promoting the practical and intelligent development of fatigue driving detection technology [7].

2. Related Work

As a crucial component of intelligent driving safety systems, driver fatigue detection has been extensively researched and gradually applied in real-world traffic scenarios. Existing driver fatigue detection methods can be broadly categorized into three categories: those based on physiological signals, those based on driving behavior, and those based on visual information. While the first two methods demonstrate high accuracy under specific conditions, they typically require additional hardware for data acquisition, resulting in high costs and potential driver discomfort [8]. In contrast, visual-based methods are gaining increasing attention due to their non-contact and convenient nature. Especially with the advancement of embedded computing power, computer vision is becoming the mainstream technology for driver fatigue detection [9].

In the field of visual fatigue detection, facial recognition and local feature analysis are key components for identifying fatigue behavior. Research has shown that fatigue is often accompanied by facial behaviors such as eye closure, increased blinking frequency, and yawning with an open mouth. Therefore, detecting the state of the eyes and mouth has become a crucial approach. Traditional image processing methods, such as feature detection based on Haar cascades or HOG+SVM, are limited by their robustness and real-time performance, and struggle to adapt to complex lighting and posture variations. In recent years, deep learning methods have become mainstream. The combination of object detection and facial landmark localization is particularly well-suited for detailed analysis of the eye and mouth regions in fatigue detection, improving detection stability and accuracy [10].

The YOLO family, a representative method in the field of object detection, strikes a good balance between real-time performance and detection accuracy. From YOLOv3 to YOLOv8, the algorithm architecture has been continuously optimized, resulting in continuous improvements in detection accuracy [11]. YOLOv8 introduces a more flexible anchor-free mechanism, an adaptive multi-scale feature fusion architecture, and a lightweight design, making it particularly suitable for real-time face detection in in-vehicle scenarios. Numerous studies have applied YOLO to driver monitoring systems, achieving rapid localization of facial regions. However, its detection accuracy for small objects such as eyes and mouths remains limited, especially in complex backgrounds and with partial occlusion, which can lead to false or missed detections [12].

To further improve the accuracy of local facial feature detection, facial landmark localization models such as FaceMesh have been widely adopted. FaceMesh, proposed by Google, can predict 468 3D facial landmarks from a single frame [13]. It offers extremely high resolution and real-time performance, making it suitable for facial behavior analysis [14]. Some research has used FaceMesh to calculate the eye aspect ratio (EAR) and mouth opening/closing ratio (MAR) to identify fatigue behaviors such as blinking and yawning. Combining YOLO with FaceMesh is a relatively new approach with great potential for exploration and application. This paper builds on this approach by designing a model architecture that integrates detection and keypoint localization, providing a more accurate and robust solution for fatigue state recognition [15].

3. Method Design

3.1. Overall System Architecture

Figure 1 shows the overall process of the fatigue driving detection system proposed in this paper. The system consists of four main modules: video input acquisition, facial region detection, key point extraction and fatigue feature analysis, and state determination and output. First, a camera captures a video stream of the driver inside the vehicle. The YOLOv8 model is used to detect the facial region. Based on this, FaceMesh extracts facial key points. The eye closure (EAR) and mouth opening/closing ratio (MAR) are then calculated. Fatigue status is then determined by combining temporal features.

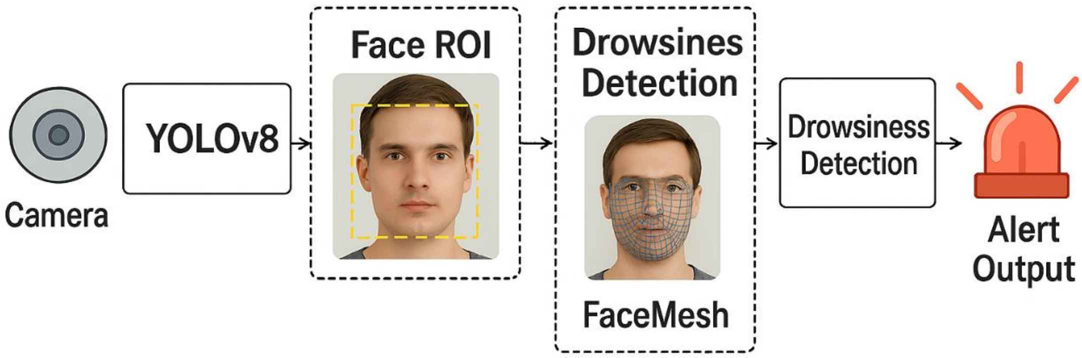

Figure 1: System overall structure diagram

3.2. YOLOv8 Face Detection Module

YOLOv8, the latest version of the object detection algorithm, utilizes an anchor-free design and an improved feature fusion module, significantly improving detection accuracy and inference speed. Its network structure primarily consists of a backbone, neck, and head. This article uses YOLOv8 for real-time face detection in vehicle-mounted scenarios.

The detection target is the driver's face area, and the network output is a bounding box (x, y, w, h) and a confidence score σ. Yolov8's advantages in this task are:

It supports lightweight deployment and is suitable for in-vehicle edge devices.

Multi-scale detection enhances response to small objects (such as eyes and mouths).

Combined with FaceMesh, it effectively narrows the search area and improves overall efficiency.

3.3. FaceMesh Key Point Extraction Module

To further accurately extract subtle motion features in the eye and mouth regions, this paper introduces Google's facemesh model, which predicts 468 3D facial key points based on a regression strategy. After YOLOv8 detects a face, it only extracts key points in the ROI region, reducing redundant computations.

As shown in Figure 2, FaceMesh can accurately mark the edge points of the eyes and mouth, facilitating subsequent feature calculations.

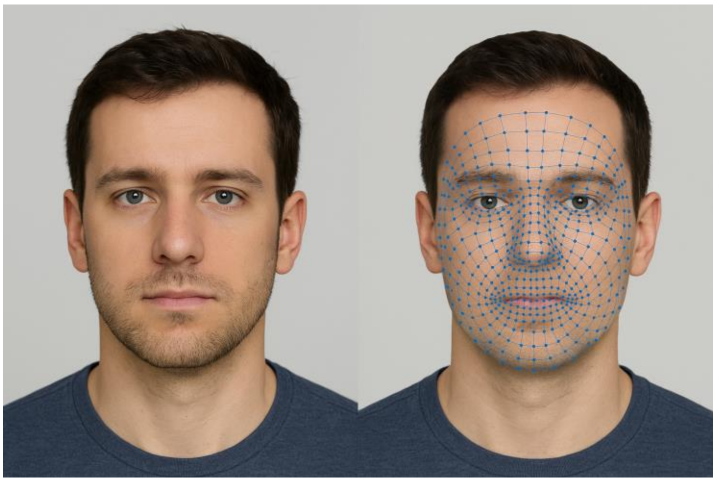

Figure 2: Schematic diagram of facial key points extracted by FaceMesh

3.4. Fatigue Behavior Feature Extraction

This study focuses on extracting two key visual features associated with fatigue:

Eye aspect Ratio (EAR): EAR is used to measure the openness of the eyes, calculated using the following formula:

$$\mathrm{EAR} = \frac{\| p_2 - p_6 \| + \| p_3 - p_5 \|}{2 \cdot \| p_1 - p_4 \|} \quad (1)$$

Where p1 to p6 represent six selected facial landmarks around the eyes extracted by FaceMesh, and ‖·‖denotes the Euclidean distance. When EAR falls below a certain threshold (e.g., 0.25) and lasts for a continuous number of frames, it can be judged as an eye closure behavior.

Mouth Aspect Ratio (MAR): MAR reflects mouth opening behavior such as yawning. It is calculated using the following formula:

$$\mathrm{MAR} = \frac{\| p_{14} - p_{18} \| + \| p_{13} - p_{19} \| + \| p_{12} - p_{20} \|}{2 \cdot \| p_{11} - p_{17} \|} \quad (2)$$

Where p11 to p20are the selected key points around the mouth region. If MAR stays above a certain threshold (e.g., 0.7) for several frames, it can be recognized as a yawning behavior.

3.5. Fatigue State Judgment Logic

Based on the extracted features, the fatigue judgment logic in this paper is defined as follows:

If the eAR value remains below a predefined threshold for more than N_1 consecutive frames, an eye-closure warning is triggered;

If the Mar value stays above a predefined threshold for more than N_2 consecutive frames, a yawn warning is triggered;

If either of the above two conditions occurs frequently (e.g., more than three times) within a given time window T, the system determines that the driver is in a state of fatigue.

A sliding window mechanism is adopted to continuously monitor the temporal variation of fatigue behaviors. This helps avoid transient misjudgments caused by momentary expressions or noise, and significantly improves the temporal robustness of the overall model.

4. Experimental Design and Result Analysis

4.1. Experimental Environment and Tool Configuration

This chapter aims to verify the effectiveness of the proposed fatigue driving detection method based on the fusion of YOLOv8 and FaceMesh. By constructing an experimental environment, selecting a public dataset, setting evaluation metrics, and comparing it with other representative methods, we comprehensively evaluate the performance of this method in fatigue feature extraction and state recognition.

The experiments were conducted on a Windows 11 operating system and an NVIDIA RTX 3060 GPU platform. The development language was Python 3.10, and the deep learning framework was PyTorch 2.0. The detection model was implemented based on Ultralytics YOLOv8. Facial landmarks were extracted using the FaceMesh module of MediaPipe. Image processing and visualization were performed using OpenCV and Matplotlib.

4.2. Dataset and Preprocessing

This study selected two public datasets for model training and evaluation: NTHU-DDD and YawDD. NTHU-DDD contains video data of multiple drivers in real-world in-car environments, covering typical fatigue behaviors such as yawning, blinking, and eye closure. The YawDD dataset focuses on open-mouth yawning, making it suitable for MAR calculation and validation. All videos were extracted frame by frame and scaled to a uniform 640×480 resolution. Human annotation tools were used to construct face box labels for training the YOLOv8 model. To enhance data diversity, image augmentation operations were also introduced, including random brightness adjustment, mirror flipping, and blur perturbation.

4.3. Evaluation Indicators and Calculation Methods

To objectively evaluate detection performance, this article uses accuracy, recall, F1 score, and frame rate (FPS) as key metrics. Accuracy measures overall recognition accuracy, recall indicates fatigue behavior coverage, F1 score comprehensively reflects the balance between precision and recall, and FPS evaluates the model's real-time processing capabilities. The mathematical expressions for these metrics are as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Inference speed is estimated by the average frame processing time, expressed in frames per second (FPS).

4.4. Experimental Results and Comparative Analysis

In fatigue signature detection experiments, the system first extracts eye and mouth key points using FaceMesh, calculates EAR and MAR values, and uses these to identify eye closure and yawning behaviors. Typical results show that when a driver enters a fatigued state, the EAR value significantly drops below 0.21, while the MAR value rises above 0.7. As shown in Figure 3, the EAR and MAR exhibit distinct temporal variations between fatigued and non-fatigued states, providing a reliable basis for fatigue assessment.

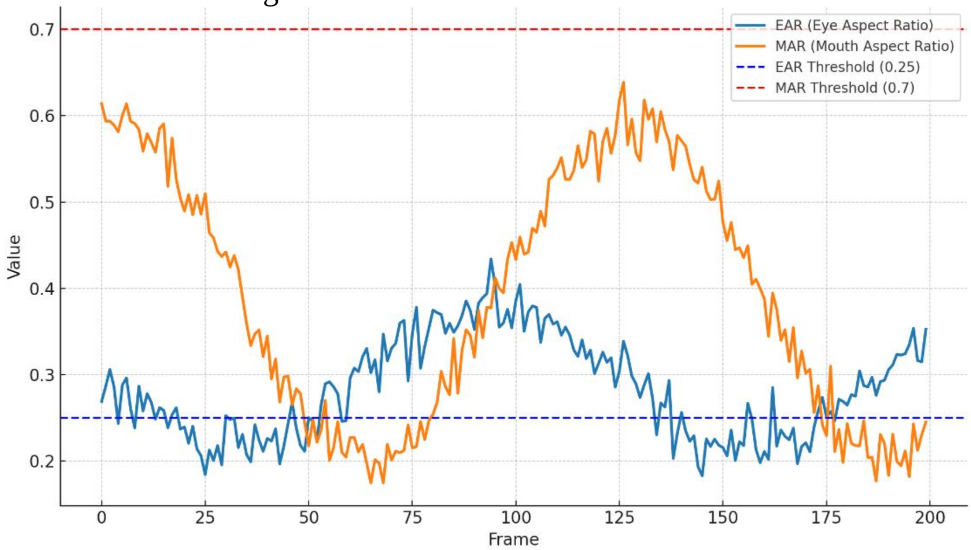

Figure 3: EAR and MAR change curves over time

Furthermore, this paper compared the performance of three model combinations on the test set: YOLOv5 combined with manual rules, YOLOv8 combined with traditional keypoint extraction methods, and the proposed yolov8+FaceMesh method. The experimental results are shown in Table 1. The proposed method significantly outperforms the comparison group in terms of accuracy, recall, and F1 value. While maintaining high detection accuracy, the inference speed remains at 39.8 FPS, meeting the requirements of real-time processing.

Table 1: Performance of three models

Model Combination	Accuracy	Recall	F1 Score	FPS
YOLOv5 + Manual Rules	82.3%	78.5%	80.1%	38.2
YOLOv8 + Traditional Keypoints	88.1%	83.9%	85.9%	41.7
YOLOv8 + FaceMesh (Proposed)	93.4%	91.6%	92.5%	39.8

To further validate the model's stability in real-world scenarios, this study designed multi-scenario robustness tests, including daytime and nighttime conditions, partial occlusion, wearing masks, and blurred images. As shown in Figure 4, the YOLOv8+FaceMesh method demonstrates good robustness in a variety of complex environments, especially in low light and light occlusion conditions, maintaining an accuracy rate of over 90% for fatigue detection.

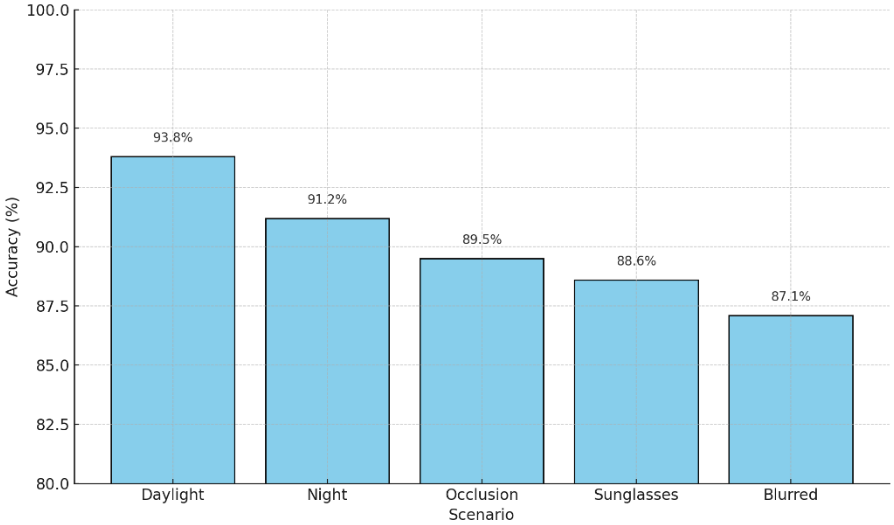

Figure 4: Histogram of model performance in each scenario

The proposed fusion method demonstrates excellent fatigue feature extraction accuracy, state recognition accuracy, and real-time performance, demonstrating its potential for application in intelligent driver monitoring systems. Compared to traditional methods, this method balances detection speed and accuracy, and exhibits good environmental adaptability, providing a practical foundation for subsequent engineering deployment.

5. Conclusion and Outlook

This paper proposes a facial fatigue detection method that integrates YOLOv8 and FaceMesh. This method aims to address the shortcomings of traditional visual methods in terms of insufficient facial region detection accuracy and weak ability to recognize subtle facial movements. By using YOLOv8 to efficiently and accurately locate facial regions and leveraging the FaceMesh model to extract detailed facial landmark information, the system achieves precise calculation of eye closure (EAR) and mouth opening (MAR). Incorporating a temporal logic judgment mechanism, the system effectively identifies driver eye closure and yawning, enabling real-time fatigue detection. Experimental results demonstrate that this method outperforms traditional model combinations in terms of accuracy, recall, and F1 score. It also demonstrates robustness and real-time performance in a variety of complex scenarios, demonstrating strong practical application value.

This research demonstrates significant innovation in its method fusion design, combining YOLOv8 with the high-density keypoint regression model FaceMesh under an anchor-free detection mechanism for the first time, achieving a balanced balance between overall detection efficiency and local feature extraction accuracy. By setting EAR and MAR thresholds and incorporating a sliding window strategy, the model can quickly identify potential fatigue states, avoiding both false positives and false negatives. Furthermore, by constructing a multi-scenario test environment, the method's adaptability to complex conditions such as occlusion, low light, and blurred images was verified, further demonstrating its scalability and potential for engineering deployment.

While this method has achieved promising results in many areas, several areas warrant further research. First, the currently used FaceMesh model relies on RGB images; future considerations include the introduction of infrared or depth cameras to enhance nighttime detection capabilities.

Second, the fatigue judgment logic is primarily based on fixed thresholds and static rules; the introduction of time series models (such as LSTM and Transformer) could enhance temporal modeling capabilities and discriminative flexibility. Finally, given the complex dynamics inherent in real-world driving environments, the construction of larger, real-world in-vehicle datasets is needed to advance this method towards higher accuracy and enhanced generalization.

References

[1] C. Li, Y. Zhu, and M. Zheng, "A multi-objective dynamic detection model in autonomous driving based on an improved YOLOv8," Alexandria Engineering Journal, vol. 122, pp. 453–464, May 2025, doi: 10.1016/j.aej.2025.03.020.

[2] A. Rahim, F. Yuan, and J. Barabady, "An ultralytics YOLOv8-based approach for road detection in snowy environments in the arctic region of Norway," Comput. Mater. Continua, vol. 83, no. 3, pp. 4411–4428, May 2025, doi: 10.32604/cmc.2025.061575.

[3] C. Zhang, X. Chen, P. Liu, B. He, W. Li, and T. Song, "Automated detection and segmentation of tunnel defects and objects using YOLOv8-CM," Tunnelling Underground Space Technol., vol. 150, p. 105857, Aug. 2024, doi: 10.1016/j.tust.2024.105857.

[4] M. Venkateswarlu and V. R. R. Chirra, "CNN-ViT: a multi-feature learning based approach for driver drowsiness detection," Array, vol. 27, p. 100425, Sep. 2025, doi: 10.1016/j.array.2025.100425.

[5] A. Rahman, M. B. H. Hriday, and R. Khan, "Computer vision-based approach to detect fatigue driving and face mask for edge computing device," Heliyon, vol. 8, no. 10, p. e11204, Oct. 2022, doi: 10.1016/j.heliyon.2022.e11204.

[6] A. Kuwahara, K. Nishikawa, R. Hirakawa, H. Kawano, and Y. Nakatoh, "Eye fatigue estimation using blink detection based on eye aspect ratio mapping(EARM)," Cognit. Rob., vol. 2, pp. 50–59, Jan. 2022, doi: 10.1016/j.cogr.2022.01.003.

[7] J. Shi and K. Wang, "Fatigue driving detection method based on time-space-frequency features of multimodal signals," Biomed. Signal Process. Control, vol. 84, p. 104744, Jul. 2023, doi: 10.1016/j.bspc.2023.104744.

[8] K. Kakhi, S. K. Jagatheesaperumal, A. Khosravi, R. Alizadehsani, and U. R. Acharya, "Fatigue monitoring using wearables and AI: trends, challenges, and future opportunities," Comput. Biol. Med., vol. 195, p. 110461, Sep. 2025, doi: 10.1016/j.compbiomed.2025.110461.

[9] Z. Cheng, "Object detection in autonomous driving scenario using YOLOv8-SimAM: a robust test in different datasets," Transportmetrica A: Transport Sci., Jun. 2025, doi: 10.1080/23249935.2025.2511818.

[10] D. Nimma, "Object detection in real-time video surveillance using attention based transformer-YOLOv8 model," Alexandria Eng. J., vol. 118, pp. 482–495, Apr. 2025, doi: 10.1016/j.aej.2025.01.032.

[11] H. Jia, Z. Xiao, and P. Ji, "Real-time fatigue driving detection system based on multi-module fusion," Comput. Graphics, vol. 108, pp. 22–33, Nov. 2022, doi: 10.1016/j.cag.2022.09.001.

[12] X. Fang, X. Yang, X. Xing, J. Wang, W. Umer, and W. Guo, "Real-time monitoring of mental fatigue of construction workers using enhanced sequential learning and timeliness," Autom. Constr., vol. 159, p. 105267, Mar. 2024, doi: 10.1016/j.autcon.2024.105267.

[13] D. Yang, B. Gao, S. Wang, and H. Xiang, "Robustness test for fouling state identification of homogeneous pressure electrodes based on confidence ellipsoids," IEICE Electron. Express, p. 22.20240283, 2025, doi: 10.1587/elex.22.20240283.

[14] X. Wang, M. Wu, C. Xu, X. Yang, and B. Cai, "State space model detection of driving fatigue considering individual differences and time cumulative effect," Int. J. Transp. Sci. Technol., vol. 13, pp. 200–212, Mar. 2024, doi: 10.1016/j.ijtst.2023.12.004.

[15] Y. Zhao, H. Zhao, and J. Shi, "YOLOv8-MAH: multi-attribute recognition model for vehicles," Pattern Recognit., vol. 167, p. 111849, Nov. 2025, doi: 10.1016/j.patcog.2025.111849.