A Learning-based Multi-Frame Visual Feature Framework for Real-Time Driver Fatigue Detection

Liang Xie[1,2], **Songlin Fan**[3,4] *
[1]School of Computer Science and Technology, Guangdong University of Technology
[2]Peking University Shenzhen Graduate School
[3]School of Computer Science, Fudan University
[4]China Mobile Shanghai ICT Co., Ltd
Liangxie@stu.pku.edu.cn,slfan@pku.edu.cn

Abstract

Driver fatigue is a significant factor contributing to road accidents, highlighting the need for reliable and accurate detection methods. In this study, we introduce a novel learning-based multi-frame visual feature framework (LMVFF) designed for precise fatigue detection. Our methodology comprises several clear and interpretable steps. Initially, facial landmarks are detected, enabling the calculation of distances between eyes, lips, and the assessment of head rotation angles based on 68 identified landmarks. Subsequently, visual features from the eye region are extracted, and an effective visual model is developed to accurately classify eye openness. Additionally, features characterizing lip movements are analyzed to detect yawning, thereby enriching fatigue detection through continuous monitoring of eye blink frequency, yawning occurrences, and head movements. Compared to conventional single-feature detection approaches, LMVFF significantly reduces instances of fatigue misidentification. Moreover, we employ various quantization and compression techniques for multiple computation stages, substantially reducing the latency of our system and achieving a real-time frame rate of 25-30 FPS for practical applications.

1 Introduction

With the rapid expansion of urban road networks, as well as the increasing mileage of train and high-speed rail lines, the frequency of traffic accidents has risen significantly. Among the various factors contributing to these accidents, fatigued driving has become one of the most concerning. The fast-paced demands of modern life have led to an alarming increase in the number of drivers experiencing fatigue behind the wheel. As shown in Fig. 1, fatigue severely (Sikander and Anwar, 2018) impairs driving reaction time, diminishes alertness,

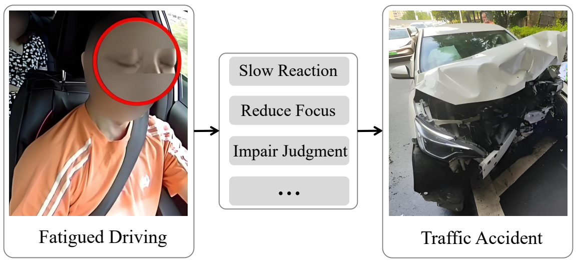


Figure 1: Extreme fatigue can cause drivers to experience slowed reactions, reduced focus, and impaired judgment, increasing the risk of accidents that may include drifting out of lanes, crashing into barriers, or failing to brake in time, leading to rear-end collisions.

and heightens the risk of poor decision-making, all of which can result in catastrophic outcomes, including serious injuries, fatalities, and significant property damage. Although intelligent road perception (Xie et al., 2022a,b) and autonomous driving (Chen et al., 2024) technologies emerge, fatigue warning remains the most effective method to prevent fatigue-related accidents (Kamti and Iqbal, 2022; Hooda et al., 2021; Yaacob et al., 2023).

Fatigued driving poses significant risks, largely because drivers often fail to recognize its symptoms until it is too late. A particularly hazardous symptom is micro-sleep, brief, involuntary episodes of attention loss lasting only a few seconds. During these lapses, vehicles traveling at high speeds may drift off the road or collide with other vehicles. Relevant research (Lian et al., 2024; He et al., 2024) has indicated that fatigued driving can be as perilous as driving under the influence of alcohol, as both conditions severely impair the ability of a driver to respond promptly and effectively to unexpected changes in traffic conditions. Given the severe risks associated with fatigued driving, researchers globally have emphasized the urgent need for effective preventive measures. To address this critical issue, various advanced detection algorithms (Li et al., 2020; Mu et al., 2017, 2024)

*Corresponding author: Songlin Fan

Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (System Demonstrations), pages 61–69

have been developed to monitor driver fatigue in real-time. These algorithms leverage physiological signals, facial analysis, and behavioral patterns to detect early signs of drowsiness and issue timely warnings, thereby reducing the likelihood of accidents. The ongoing enhancement and deployment of these technological solutions are vital in mitigating the dangers posed by fatigued driving and improving road safety for all drivers.

Existing fatigued driving detection methods can be classified into two categories, i.e., traditional and learning-based methods. Traditional methods rely on rule-based approaches, statistical analysis, and signal processing techniques. Since handcrafted features from traditional methods (Kreucher et al., 1998; Wang et al., 2018) have limited representation capability, they often fail to capture complex fatigue patterns accurately and may exhibit poor adaptability under varying real-world driving conditions. Due to their powerful modeling capability, learning-based approaches (Zhuang et al., 2020; Jia et al., 2021; Lambay et al., 2024; Hosseini and Horbach, 2023) have become increasingly preferred for developing advanced driver fatigue detection systems, presenting promising opportunities. However, existing learning-based methods often suffer from suffer from narrow feature coverage, limited interpretability, and excessively bulky models.

To address these issues, we propose a learning-based multi-frame visual feature framework named LMVFF, which integrates multi-dimensional feature modeling within a lightweight design. First, the framework performs facial landmark detection and calculates distances between eyes, distances within the lip region, and head rotation angles derived from 68 facial landmarks to preliminarily assess fatigue. It then extracts visual features from the eye region to construct a visual model that determines eye openness. Subsequently, it integrates open and closed lip-region features to identify yawning behaviors, evaluating driver fatigue based on blinking and yawning frequencies to minimize misjudgments caused by single-feature reliance. Comprehensive experiments on various datasets demonstrate that the proposed method achieves high accuracy in fatigue detection.

2 Related Work

2.1 Traditional Methods

There is a wide range of traditional fatigue detection methods that leverage physical signals. For instance, some methods utilize Electroencephalogram (EEG) (Mu et al., 2017; Wang et al., 2018) signals to differentiate between wakefulness and sleep states. However, EEG signals are prone to noise interference and can be challenging to collect. Electrooculogram (EOG) (Reilly and Lee, 2010) signals are more accessible to gather and less sensitive to minor disturbances, though they still require head-mounted devices for data acquisition. Behavioral feature-based methods (Tuncer et al., 2021) utilize remote video cameras to capture real-time images, monitoring parameters such as eyelid movement, facial orientation, and gaze direction. In addition, lane detection methods (Salvati et al., 2021) employ cameras to monitor the road ahead, assessing vehicle control by analyzing lane markings.

These state-based methods (Albadawi et al., 2022) do not require physical contact with the driver and demand minimal equipment, which enhances their practicality and ease of implementation. Nonetheless, varying road conditions and differences in vehicle models (Kreucher et al., 1998; Fu et al., 2016; Kong et al., 2015) can constrain their effectiveness. Meanwhile, Mu *et al.* (Mu et al., 2024) propose a hierarchical transformer-based fatigue detection system using ECG and HRV features, enhancing prediction accuracy with multi-scale representations, efficient attention, and adaptive feature fusion. Hu *et al.* (Hu et al., 2024) introduce a spatio-temporal fusion network (STFN-BRPS) using multi-scale convolution, dynamic graph convolution, and feature fusion modules for enhanced driver fatigue detection using EEG signals, improving accuracy and robustness.

2.2 Learning-based Methods

Manually designed features often have poor discriminative ability, and relying on a single visual feature can be problematic. For example, glare from glasses can prevent the camera from capturing eye movements, the degree of eye openness varies between individuals, and irregular head movements can cause false alarms (El-Nabi et al., 2024; Zhang et al., 2023). Therefore, there is a need for efficient and accurate strategies for fatigue detection. With the advancement of deep learning, many fatigue-driving detection algorithms have shown some effectiveness and progress, such as approaches (Li et al., 2017; Zhuang et al., 2020) primarily include automatic fatigue detection based on the facial information of drivers and eye fatigue detection based on pupil and iris segmentation. These methods yield poor detection results and have low accuracy. Methods (You et al., 2020; Jia et al., 2021) employ strategies such as facial feature integration, facial motion entropy, and multi-index fusion. Lambay *et al.* (Lambay et al., 2024) review machine learning techniques for detecting and monitoring physical fatigue in manufacturing and human–robot collaboration, analyzing detection complexity, influencing factors, feature selection, and monitoring challenges.

Cheng *et al.* (Cheng et al., 2024) introduce a multi-feature fusion fatigue detection algorithm using dual cameras and a lightweight CNN to analyze facial and lane departure features, enhancing accuracy and real-time performance for effective driver fatigue warning systems. Wu *et al.* (Wu et al., 2025) propose a multi-dimensional adaptive transformer network for driver fatigue detection, leveraging weighted feature extraction across EEG dimensions to enhance accuracy, structural information retrieval, and model generalization. Zhou *et al.* (Zhou et al., 2024) explore YOLOv8-based fatigued driving detection, analyzing methods, datasets, and global research to enhance prevention, reduce accidents, and improve road safety. Li *et al.* (Li et al., 2024) propose NAS-optimized lightweight CNN models for real-time driving fatigue detection using multichannel EEG, ensuring high performance and efficiency for deployment in intelligent vehicles. However, due to the complexity of the models designed, they are not suitable for real-time detection scenarios.

3 Architecture

The proposed learning-based multi-frame visual feature framework (LMVFF) for fatigued driving detection is illustrated in Fig. 2. Since our multi-frame version (see Fig. 2 (a)) is built on the single-frame one (see Fig. 2 (b)), this section will detail the single-frame method and then transition the multi-frame version.

Face Detection. Face images collected from various scenarios in train and vehicle cabins are initially pre-processed. A face detection model is trained using Libfacedetection (Wu et al., 2023; Feng et al., 2022), achieving real-time performance at approximately 150 FPS. To minimize false positives, we also integrate the Onet network from MTCNN (Xiang and Zhu, 2017), effectively balancing speed and accuracy.

Face Landmark Detection. Leveraging publicly available and self-collected datasets, we annotate images with 68 face landmarks. A landmark detection model is developed using MobileNetv2 (Sandler et al., 2018) combined with depthwise separable convolutions (Chollet, 2017) and a weighted loss function to enhance precision in landmark localization. MobileNetV2 is a highly efficient convolutional neural network architecture designed for resource-constrained devices. Its lightweight design makes it particularly well-suited for real-time tasks such as face landmark detection, where speed and accuracy are critical.

Face Alignment and Feature Extraction. After landmark detection, the face image is aligned by computing a similarity transformation matrix (Kortli et al., 2020) based on these facial landmarks:

$$\mathbf{T} = \arg\min_{\mathbf{T}} \sum_{i=1}^{N} |\mathbf{x}'_i - \mathbf{T}\mathbf{x}_i|^2, \quad (1)$$

where $\mathbf{x}_i$ and $\mathbf{x}'_i$ are the original landmarks and target landmarks, respectively. $\mathbf{T}$ is the similarity transformation matrix. Subsequently, a feature extraction model based on ArcFace (Deng et al., 2019) is trained, reducing intra-class variance and increasing inter-class variance, thereby yielding distinctive facial features.

Face Feature Matching. A driver face recognition database, encompassing diverse scenarios and orientations, is constructed. Cosine similarity $S(\cdot)$ between the face feature vectors (Chen et al., 2021) is utilized for face matching:

$$S(\mathbf{f}_a, \mathbf{f}_b) = \frac{\mathbf{f}_a \cdot \mathbf{f}_b}{|\mathbf{f}_a||\mathbf{f}_b|}, \quad (2)$$

where $\mathbf{f}_a$ and $\mathbf{f}_b$ are the current face features and the recorded face features, respectively. Face matching results enable driver verification and attendance logging by recording initial login timestamps.

Eye State Recognition. Then, aligned faces are cropped to extract left and right eye regions, categorized into open and closed eyes. A MobileNetv2 classification model is trained to classify these eye states, represented numerically (The predicted result 0 for open, 1 for closed).

Mouth State Recognition. Similarly, the mouth region is isolated from aligned faces, categorized into open and closed states, and used to train a MobileNetv2 classification model. The states are numerically encoded as 0 for a closed mouth and 1 for an open mouth.

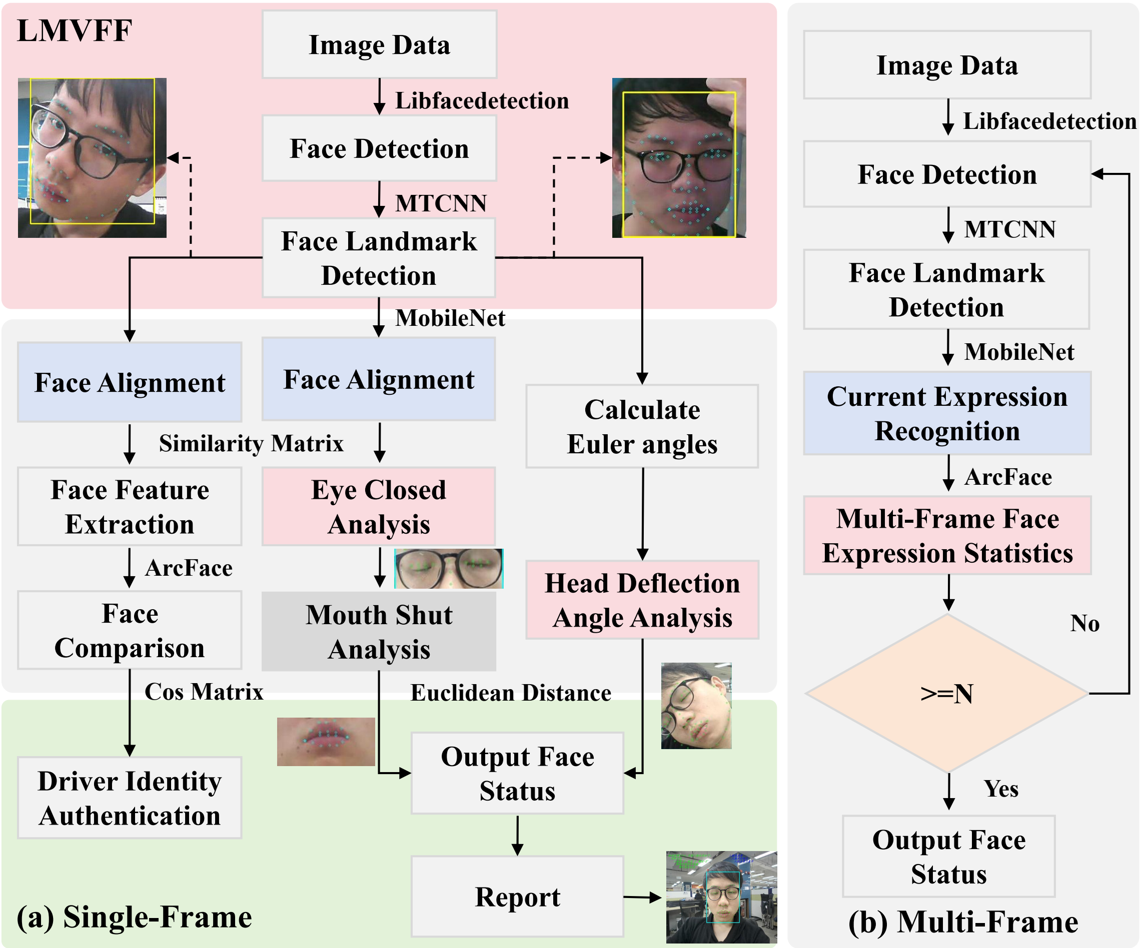


Figure 2: The framework of fatigued driving detection. It can be divided into two parts: single-frame detection and multi-frame detection. Single-frame detection makes judgments based on the results of merging different features, while multi-frame detection analyzes continuous feature information.

Head Pose Detection. Using a subset of 14 critical face landmarks from the detected 68 ones, Euler angles (Yaw, Pitch, Roll) are calculated (Slabaugh, 1999). Thresholds of $\pm 60°$(Yaw), $\pm 30°$ (Pitch), and $\pm 30°$ (Roll) are applied to assess head orientation, thereby identifying inattentiveness.

Eye State Analysis. Landmark-based distance ratios around the pupil are computed and compared against predefined thresholds to validate eye states. This landmark-based method complements the MobileNetv2 model predictions, significantly reducing false positives.

Mouth State Analysis. Meanwhile, we calculate the distance ratios of the upper and lower lips to the mouth center, compare them with a set threshold, and combine these results with the classification model to accurately determine mouth state. These steps allow single-frame face image to be analyzed for fatigued driving detection, providing real-time analysis and generating detailed monitoring reports in the end.

Multi-Frame Analysis. Finally, based on the single-frame driver fatigue detection method introduced above, for long video sequences, we conduct statistical analysis across multiple frames, as shown in Fig. 2 (b). Each frame is individually analyzed and the results are stored in a buffer queue. If more than five frames indicate signs of fatigue, an alert is triggered, and subsequent frames are continuously analyzed.

4 Experiment

Training and Testing Data: We utilize multiple publicly available facial landmark datasets, such as 300W (Sagonas et al., 2016) and AFLW (Koestinger et al., 2011), for pretraining. Additionally, we collect a substantial amount of frontal and profile face data as the test set. The test data is divided into five groups (as shown in Table 1), each focusing on different facial angles to ensure the model's generalization across various viewpoints and angles.

Table 1: Test results in different scenarios. We collect five sets of complex frontal and profile face datasets from different regions and used the NME (Lai et al., 2019) as the evaluation metric, with the inter-point distance as the normalization factor. We then calculate the NME values for different facial regions, such as the eyes, nose, mouth, contour, and eyebrows. Finally, we conduct experiments using different trained versions of the MobileNetv2 (Sandler et al., 2018) and VGG (Simonyan and Zisserman, 2014) models as landmark prediction models.

	Dataset	Face	MobileNetv2_L55_64	MobileNetv2_L30_64	VGG_L7_224	VGG_L7_64
Center	Area 1 Positive Face (1296)	Eye	3.571	3.257	2.793	5.159
		Nose	3.450	3.379	2.820	5.047
		Brow	5.706	5.152	4.783	7.113
		Mouth	3.486	3.497	3.299	6.085
		Contour	5.038	4.888	4.563	8.720
	Area 2 Profile Face (1263)	Eye	4.692	4.711	5.844	6.810
		Nose	4.370	4.524	5.466	6.603
		Brow	6.349	6.363	7.819	9.418
		Mouth	4.370	4.664	6.200	7.313
		Contour	6.441	7.236	8.458	10.747
	Area 3 Positive Face (583)	Eye	4.201	3.895	3.283	5.764
		Nose	4.269	3.936	3.314	5.339
		Brow	5.816	5.739	4.737	7.761
		Mouth	3.988	3.624	3.272	5.486
		Contour	6.241	5.932	5.387	8.762
	Area 4 Profile Face (456)	Eye	7.495	6.675	8.562	10.186
		Nose	5.792	5.577	8.028	7.590
		Brow	7.246	6.529	8.209	10.462
		Mouth	5.607	5.236	7.878	7.858
		Contour	9.267	8.491	10.423	11.486
	Area 5 Positive Face (1049)	Eye	10.084	9.868	22.014	17.350
		Nose	10.959	10.892	28.434	18.877
		Brow	12.463	11.682	24.464	20.908
		Mouth	11.775	11.161	29.414	17.030
		Contour	19.161	18.123	27.629	21.091

Evaluation Metrics: For the landmark detection task, we employ NME (Normalized Mean Error (Lai et al., 2019)) as the evaluation metric, which measures the average distance between the detected and ground-truth landmark to assess the facial recognition accuracy of the model. We adopt MSE (Mean Squared Error (Wu and Ji, 2019)) as the loss function in training to optimize the landmark prediction performance.

Meanwhile, for face detection task, we employ MTCNN (Xiang and Zhu, 2017) network, while landmark detection is performed using different versions of MobileNetV2 and VGG models, with tests conducted on various network depths (as shown in Table 2). Furthermore, to improve the accuracy of frontal and profile face classification, we apply different classification strategies to a meticulously designed dataset and measured accuracy as the proportion of correctly classified samples relative to the total dataset.

4.1 Landmark Analysis

In this study, we utilize different trained MobileNetv2 and VGG models as landmark prediction models. To more precisely evaluate the landmark error in various facial regions, we divide the 68 facial landmark (indexed 1-68 as shown in Fig. 3) into different areas for testing and calculate the normalized mean error values, using inter-point distance as the normalization factor in NME. The

Table 2: The classification accuracy in different scenarios. We use different combinations to determine whether a face is front-facing. "Crosswire" selects five points from the 68 predicted by the landmark model and filters out the proportion of small angle and large angle faces using a five-point cross-line method. "Midcourt" calculates classification accuracy based on the median line, while "Crosswire + Midcourt" combines both methods to determine overall accuracy. "Euler Angle" classifies faces using Euler equation, whereas Contour Area employs contour area as a filtering criterion. "Pupil + Dilateral + Midcourt" integrates pupil distance, bilateral positioning, and the median line for classification, whereas "Pupil + Dilateral + 27 Point" incorporates pupil distance, bilateral positioning, and the distances from 27 points to both pupils.

Various Combination	Angle = $0° \sim 60°$		Angle = $60° \sim 90°$	
Crosswire	Small	Large	Small	Large
MTCNN (Xiang and Zhu, 2017)	0.869	0.131	0.143	0.857
MobileNetv2 (Sandler et al., 2018)	0.861	0.139	0.185	0.815
Midcourt	Small	Large	Small	Large
MTCNN (Xiang and Zhu, 2017)	0.928	0.072	0.224	0.776
MobileNetv2 (Sandler et al., 2018)	0.934	0.066	0.124	0.876
Crosswire + Midcourt	Small	Large	Small	Large
MTCNN (Xiang and Zhu, 2017)	0.867	0.133	0.118	0.882
MobileNetv2 (Sandler et al., 2018)	0.858	0.142	0.085	0.915
Euler Angle	Small	Large	Small	Large
MobileNetv2 (Sandler et al., 2018)	0.9	0.1	0.116	0.884
Contour Area	Small	Large	Small	Large
MobileNetv2 (Sandler et al., 2018)	0.83	0.17	0.84	0.16
Pupil distance + Bilateral + Midcourt	Small	Large	Small	Large
MobileNetv2 (Sandler et al., 2018)	0.968	0.032	0.021	0.979
Pupil distance + Bilateral + 27 Points	Small	Large	Small	Large
MobileNetv2 (Sandler et al., 2018)	0.969	0.031	0.023	0.977

specific divisions are as follows: the entire face (68 landmarks: 1-68), the non-contour region (51 landmarks: 18-68), the facial contour region (17 landmarks: 1-17), the eyebrow region (10 landmarks: 18-27), the nose region (9 landmarks: 28-36), the eye region (12 landmarks: 37-48), and the mouth region (20 landmarks: 49-68). Additionally, we compute the NME values using five different datasets collected independently. These values can serve as indicators for detecting abnormalities in the mouth, eyes, and facial angles. Finally, as shown in Table 1, the models we employ exhibit high accuracy in landmark prediction, and the calculated NME values are also precise. These findings provide a solid foundation for detecting abnormal facial states in drivers.

4.2 Different Strategies Analysis

We need to compute the facial landmarks for different algorithms to facilitate subsequent comparative analysis. Specifically, the MTCNN detector employs a five-point landmark scheme, processing all test datas through the MTCNN detector to extract and save facial bounding boxes and landmark information as a baseline for testing and comparison. In contrast, the MobileNetv2 model adopts a 68-point landmark approach, where test data is processed using an libfacedetection detector to generate facial bounding boxes, followed by MobileNetv2 predictions to obtain 68 landmark coordinates.

After acquiring the 68 points, we select 7 landmarks for Euler angle conversion and synthesize five landmarks from these 68 points for compari-

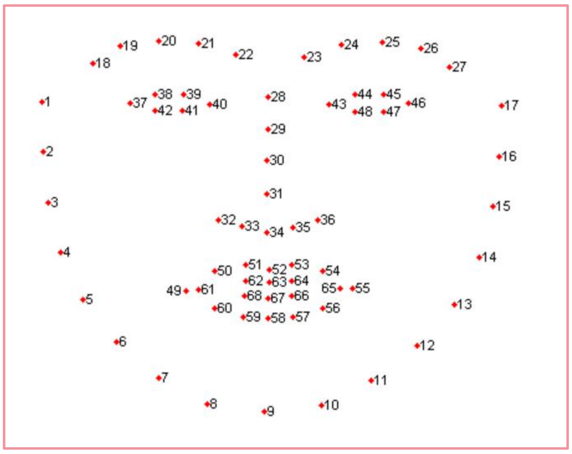
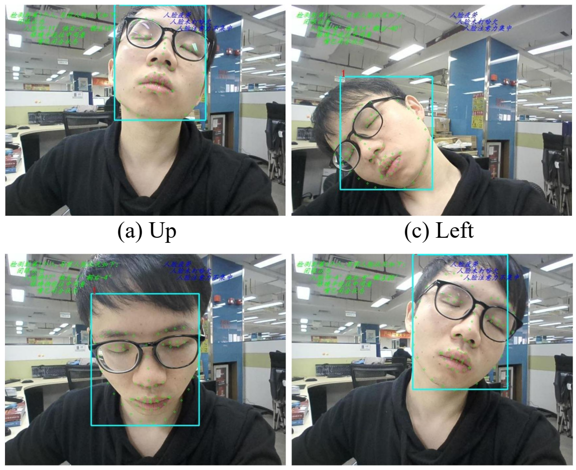


Figure 3: Our selected 68 facial landmark coordinate positions in the testing.

son with MTCNN's five-point results using a side-face ratio filtering scheme. Meanwhile, in Table 2, "Small" denotes a small angle range, while "Large" indicates a large angle range. The experiment designed various ratio judging strategies: the mean of left eye landmarks (37, 38, 40, 41) is computed to synthesize the left eyeball point, and the mean of right eye landmarks (43, 44, 46, 47) is used to synthesize the right eyeball point, combined with points 30, 48, and 54 to form a five-point set, namely left eye (0), right eye (1), nose (2), left mouth corner (3), and right mouth corner (4).

Next, we design multiple classification schemes to evaluate their performance based on these landmark judgment strategies, including "Crosswire" (five-point crossline), "Midcourt" (median line), "Crosswire + Midcourt" (a combination of both methods), "Euler Angle" (classification values computed using Euler angles), "Contour Area" (based on contour area), "Pupil + Dilateral + Midcourt" (a combination of pupil distance, bilateral mid-center, and median line), and "Pupil + Dilateral + 27 Point" (a combination of pupil distance, bilateral mid-center, and the distance from point 27 to the left and right pupils). Through experimental comparison, we found that the combination of "Pupil + Dilateral + 27 Point" performed the best.

To further optimize the filtering process, we establish the following judgment thresholds: filtering is applied when the pupil distance is less than 20 pixels. Judging occurs if the bilateral midline center of the five landmarks is on the nose side (either left or right), and filtering is triggered if the excitation ratio of the distance from point 27 to the left and right pupils is less than 0.6. Additionally, for the MobileNetv2 68-point classification



(a) Up (c) Left

(b) Down (d) Right

Figure 4: The visualization results of fatigue detection. The states of head shaking (left to right and up to down) and eye closing are detected from top to bottom, with the yellow and blue fonts displaying our test results. For more demonstrations, please refer to our video demo.

scheme in the angle, we utilize the Pitch and Roll angles from Euler angles for judgment, with recommended thresholds set at Pitch $\leq -16°$ or Pitch $\geq 16°$, and Roll $\leq -35°$ or Roll $\geq 35°$. Note that these angles may deviate from actual values and can be adjusted according to specific needs.

The visualization results of fatigued driving detection are shown in Fig. 4. Our method can accurately detect various head rotation angles, mouth opening, and eye closing actions, allowing for precise determination of driver fatigue.

5 Conclusion

In conclusion, we present a learning-based approach for detecting driver fatigue through multi-frame facial analysis, incorporating multiple optimization strategies to enhance detection accuracy and system latency. We employ MobileNetv2 as the backbone for landmark detection and apply enhanced facial feature weights to improve eye landmark perception accuracy. Based on these perceived landmarks, we crop the eye region and introduce an effective eye state classification model to accurately identify closed eyes. Similarly, the mouth region is cropped using facial landmarks, enabling a new mouth state classification model that significantly boosts yawning detection accuracy. Lastly, we leverage Euler angles to determine head orientation. Collectively, the features from these techniques enable robust identification of potential driver fatigue.

References

Yaman Albadawi, Maen Takruri, and Mohammed Awad. 2022. A Review of Recent Developments in Driver Drowsiness Detection Systems. *Sensors*, 22(5):2069.

Jiawei Chen, Zhengwei Guo, and Junlin Hu. 2021. Ring-regularized Cosine Similarity Learning for Fine-grained Face Verification. *Pattern Recognition Letters*, 148:68–74.

Li Chen, Penghao Wu, Kashyap Chitta, Bernhard Jaeger, Andreas Geiger, and Hongyang Li. 2024. End-to-end autonomous driving: Challenges and frontiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Wangfeng Cheng, Xuanyao Wang, and Bangguo Mao. 2024. A Multi-feature Fusion Algorithm for Driver Fatigue Detection Based on a Lightweight Convolutional Neural Network. *The Visual Computer*, 40(4):2419–2441.

François Chollet. 2017. Xception: Deep Learning with Depthwise Separable Convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.

Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. 2019. Arcface: Additive Angular Margin Loss for Deep Face Recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4690–4699.

Samy Abd El-Nabi, Walid El-Shafai, El-Sayed M El-Rabaie, Khalil F Ramadan, Fathi E Abd El-Samie, and Saeed Mohsen. 2024. Machine Learning and Deep Learning Techniques for Driver Fatigue and Drowsiness Detection: A Review. *Multimedia Tools and Applications*, 83(3):9441–9477.

Yuantao Feng, Shiqi Yu, Hanyang Peng, and Yan-Ran Li. 2022. Detect Faces Efficiently: A Survey and Evaluations. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(1):1–18.

Rongrong Fu, Hong Wang, and Wenbo Zhao. 2016. Dynamic Driver Fatigue Detection Using Hidden Markov Model in Real Driving Condition. *Expert Systems with Applications*, 63:397–411.

Chen He, Pengpeng Xu, Xin Pei, Qianfang Wang, Yun Yue, and Chunyang Han. 2024. Fatigue at the Wheel: A Non-visual Approach to Truck Driver Fatigue Detection by Multi-feature Fusion. *Accident Analysis & Prevention*, 199:107511.

Rohit Hooda, Vedant Joshi, and Manan Shah. 2021. A Comprehensive Review of Approaches to Detect Fatigue Using Machine Learning Techniques. *Chronic Diseases and Translational Medicine*.

Mohammad Hosseini and Serge PJM Horbach. 2023. Fighting Reviewer Fatigue or Amplifying Bias? Considerations and Recommendations for Use of ChatGPT and Other Large Language Models in Scholarly Peer Review. *Research integrity and peer review*, 8(1):4.

Fo Hu, Lekai Zhang, Xusheng Yang, and Wen-An Zhang. 2024. EEG-Based Driver Fatigue Detection using Spatio-Temporal Fusion Network with Brain Region Partitioning Strategy. *IEEE Transactions on Intelligent Transportation Systems*.

Huijie Jia, Zhongjun Xiao, and Peng Ji. 2021. Fatigue Driving Detection Based on Deep Learning and Multi-Index Fusion. *IEEE Access*, 9:147054–147062.

Mukesh Kumar Kamti and Rauf Iqbal. 2022. Evolution of Driver Fatigue Detection Techniques—A Review from 2007 to 2021. *Transportation research record*, 2676(12):485–507.

Martin Koestinger, Paul Wohlhart, Peter M Roth, and Horst Bischof. 2011. Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization. In *IEEE international conference on computer vision workshops*, pages 2144–2151. IEEE.

Wanzeng Kong, Lingxiao Zhou, Yizhi Wang, Jianhai Zhang, Jianhui Liu, and Shenyong Gao. 2015. A System of Driving Fatigue Detection Based on Machine Vision and its Application on Smart Device. *Journal of Sensors*, 2015(1):548602.

Yassin Kortli, Maher Jridi, Ayman Al Falou, and Mohamed Atri. 2020. Face Recognition Systems: A Survey. *Sensors*, 20(2):342.

Chris Kreucher, Sridhar Lakshmanan, and Karl Kluge. 1998. A Driver Warning System Based on the LOIS Lane Detection Algorithm. In *Proceedings of IEEE international conference on intelligent vehicles*, volume 1, pages 17–22. Stuttgart, Germany.

Shenqi Lai, Zhenhua Chai, Shengxi Li, Huanhuan Meng, Mengzhao Yang, and Xiaoming Wei. 2019. Enhanced Normalized Mean Error loss for Robust Facial Landmark detection. In *BMVC*, page 111.

Arsalan Lambay, Ying Liu, Phillip L Morgan, and Ze Ji. 2024. Machine Learning Assisted Human Fatigue Detection, Monitoring, and Recovery. *Digital Engineering*, page 100004.

Kening Li, Yunbo Gong, and Ziliang Ren. 2020. A Fatigue Driving Detection Algorithm Based on Facial Multi-Feature Fusion. *IEEE Access*, 8:101244–101259.

Qingqing Li, Zhirui Luo, Ruobin Qi, and Jun Zheng. 2024. Automatic Searching of Lightweight and High-performing CNN Architectures for EEG-based Driving Fatigue Detection. *IEEE Transactions on Instrumentation and Measurement*.

Zuojin Li, Liukui Chen, Jun Peng, and Ying Wu. 2017. Automatic Detection of Driver Fatigue Using Driving Operation Information for Transportation Safety. *Sensors*, 17(6):1212.

Zequan Lian, Tao Xu, Zhen Yuan, Junhua Li, Nitish Thakor, and Hongtao Wang. 2024. Driving Fatigue Detection Based on Hybrid Electroencephalography and Eye Tracking. *IEEE Journal of Biomedical and Health Informatics*.

Siqi Mu, Shiwei Liao, Kuan Tao, and Yanfei Shen. 2024. Intelligent Fatigue Detection Based on Hierarchical Multi-scale ECG Representations and HRV Measures. *Biomedical Signal Processing and Control*, 92:106127.

Zhendong Mu, Jianfeng Hu, and Jianliang Min. 2017. Driver Fatigue Detection System Using Electroencephalography Signals Based on Combined Entropy Features. *Applied Sciences*, 7(2):150.

Richard B Reilly and T Clive Lee. 2010. Electrograms (ecg, eeg, emg, eog). *Technology and Health Care*, 18(6):443–458.

Christos Sagonas, Epameinondas Antonakos, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 2016. 300 Faces in-the-wild Challenge: Database and Results. *Image and vision computing*, 47:3–18.

Luca Salvati, Matteo d'Amore, Anita Fiorentino, Arcangelo Pellegrino, Pasquale Sena, and Francesco Villecco. 2021. On-road Detection of Driver Fatigue and Drowsiness During Medium-Distance Journeys. *Entropy*, 23(2):135.

Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. 2018. Mobilenetv2: Inverted Residuals and Linear Bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520.

Gulbadan Sikander and Shahzad Anwar. 2018. Driver fatigue detection systems: A review. *IEEE Transactions on Intelligent Transportation Systems*, 20(6):2339–2352.

Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-scale Image Recognition. *arXiv preprint arXiv:1409.1556*.

Gregory G Slabaugh. 1999. Computing Euler Angles from a Rotation Matrix. *Retrieved on August*, 6(2000):39–63.

Turker Tuncer, Sengul Dogan, and Abdulhamit Subasi. 2021. EEG-based Driving Fatigue Detection Using Multilevel Feature Extraction and Iterative Hybrid Feature Selection. *Biomedical Signal Processing and Control*, 68:102591.

Hongtao Wang, Andrei Dragomir, Nida Itrat Abbasi, Junhua Li, Nitish V Thakor, and Anastasios Bezerianos. 2018. A Novel Real-time Driving Fatigue Detection System Based on Wireless Dry EEG. *Cognitive neurodynamics*, 12(4):365–376.

Dingming Wu, Liu Deng, Quanping Lu, and Shihong Liu. 2025. A Multidimensional Adaptive Transformer Network for Fatigue Detection. *Cognitive Neurodynamics*, 19(1):43.

Wei Wu, Hanyang Peng, and Shiqi Yu. 2023. YuNet: A Tiny Millisecond-level Face Detector. *Machine Intelligence Research*, pages 1–10.

Yue Wu and Qiang Ji. 2019. Facial Landmark Detection: A Literature Survey. *International Journal of Computer Vision*, 127(2):115–142.

Jia Xiang and Gengming Zhu. 2017. Joint Face Detection and Facial Expression Recognition with MTCNN. In *2017 4th international conference on information science and control engineering*, pages 424–427. IEEE.

Liang Xie, MengHao Hu, and XinBei Bai. 2022a. Online Improved Vehicle Tracking Accuracy via Unsupervised Route Generation. In *IEEE 34th International Conference on Tools with Artificial Intelligence*, pages 788–792. IEEE.

Liang Xie, MengHao Hu, and XinBei Bai. 2022b. Towards Hardware-Friendly and Robust Facial Landmark Detection Method. In *International Conference on Neural Information Processing*, pages 432–444. Springer.

Hamwira Yaacob, Farhad Hossain, Sharunizam Shari, Smith K Khare, Chui Ping Ooi, and U Rajendra Acharya. 2023. Application of Artificial Intelligence Techniques for Brain–Computer Interface in Mental Fatigue Detection: A Systematic Review (2011–2022). *IEEe Access*, 11:74736–74758.

Feng You, Yunbo Gong, Haiqing Tu, Jianzhong Liang, and Haiwei Wang. 2020. A Fatigue Driving Detection Algorithm Based on Facial Motion Information Entropy. *Journal of advanced transportation*, 2020(1):8851485.

Hui Zhang, Dingan Ni, Naikan Ding, Yifan Sun, Qi Zhang, and Xin Li. 2023. Structural Analysis of Driver Fatigue Behavior: A Systematic Review. *Transportation research interdisciplinary perspectives*, 21:100865.

Chang Zhou, Yang Zhao, Shaobo Liu, Yi Zhao, Xingchen Li, and Chiyu Cheng. 2024. Research on Driver Facial Fatigue Detection Based on Yolov8 Model. In *5th International Conference on Information Science, Parallel and Distributed Systems*, pages 282–285. IEEE.

Qianyang Zhuang, Zhang Kehua, Jiayi Wang, and Qianqian Chen. 2020. Driver Fatigue Detection Method Based on Eye States With Pupil and Iris Segmentation. *Ieee Access*, 8:173440–173449.