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

A Driver Emotion Monitoring System Based on Computer Vision Technique

Zekai Yan*

Graduate School of Arts and Sciences, Columbia University, New York 10027, USA

Author to whom correspondence should be addressed.

Abstract: Abnormal driving behavior includes driving distraction, fatigue, road anger, phone use, and an exceptionally happy mood. Detecting abnormal driving behavior in advance can avoid traffic accidents and reduce the risk of traffic conflicts. Traditional methods of detecting abnormal driving behavior include using wearable devices to monitor blood pressure, pulse, heart rate, blood oxygen, and other vital signs, and using eye trackers to monitor eye activity (such as eye closure, blinking frequency, etc.) to estimate whether the driver is excited, anxious, or distracted. Traditional monitoring methods can detect abnormal driving behavior to a certain extent, but they will affect the driver's normal driving state, thereby introducing additional driving risks. This research uses the combined method of support vector machine and dlib algorithm to extract 68 facial feature points from the human face, and uses an SVM model as a strong classifier to classify different abnormal driving statuses. The combined method reaches high accuracy in detecting road anger and fatigue status and can be used in an intelligent vehicle cabin to improve the driving safety level.

Keywords: Abnormal driving behavior; Support vector machine (SVM); Dlib algorithm (facial feature); Driving safety

Online publication: August 7, 2025

1. Introduction

According to statistics, more than 30% of road traffic accidents are related to fatigue driving and road rage driving, about 55–70% of major traffic accidents are related to fatigue and road rage driving, and about 46% of small and medium-sized traffic accidents are related to distracted driving (including mobile phone use, smoking, etc.) [1]. Abnormal driving behavior seriously affects the driver's alertness, adaptability, and safe driving ability, and is prone to serious traffic accidents with mass casualties [2]. How to automatically identify bad driving status through computer vision methods is a technical problem that needs to be solved in the fields of traffic engineering and artificial intelligence. Solving this technical difficulty will greatly promote the overall improvement of traffic driving safety.

Several studies have explored machine learning approaches for detecting driver emotions, particularly anger, with varying methodologies and performance outcomes. Gao proposed a real-time non-intrusive monitoring system using linear SVMs to classify driver anger and disgust [3]. Their system achieved 86.2% accuracy in in-car scenarios, demonstrating the feasibility of lightweight models for emotion detection. Wan leveraged Least Squares Support Vector Machine (LSSVM) to identify driving anger states based on 13 multimodal features, including skin conductance (SC), EEG signals, and vehicle speed (SP) [4]. Their model attained 82.20% accuracy, highlighting the utility of physiological and vehicular data for emotion recognition. Huang combined SVM classifiers with CNN for emotion detection by integrating facial expressions and EEG signals [5]. However, their approach yielded a lower accuracy (<70%), suggesting challenges in fusing heterogeneous data sources effectively.

Fatigued driving, typically caused by insufficient sleep or improper cognitive load during driving [6], poses significant safety risks that are often overlooked and challenging to quantify [7]. Epidemiological studies demonstrate that fatigued driving contributes to approximately 20% of fatal collisions worldwide [8]. As China's motorization rate continues to grow at an annual rate of 7.3% [9], the dangers of fatigued driving may escalate further.

Neurocognitive research reveals that fatigue impairs driving performance through multiple mechanisms: slower reaction times by 30–50% according to Dawson [6], increased distraction propensity [10], reduced vigilance measured by EEG spectral analysis [11], and riskier driving behaviors as quantified by lane deviation metrics [12]. However, current countermeasures face three major challenges: (1) the lack of operationalized fatigue definitions [13], (2) technological limitations in detection systems (with only 68% accuracy in current commercial devices per a 2023 IEEE review), and (3) implementation barriers for legislative solutions [7].

While numerous studies have examined anger expressions through whole-face analysis, individual facial variations can significantly influence judgment accuracy. To address this, our study employs 68 lightweight facial landmarks to isolate key features (eyes, eyebrows, and mouth) that demonstrate the strongest correlation with anger and fatigue [14]. Support Vector Machine (SVM) with radial kernel was built to process these extracted features to calculate anger and fatigue probability scores for each facial component. Using probabilistic predictions, we established a 0.8 decision threshold to minimize false positives (misclassifying neutral expressions as anger), thereby reducing potential driver annoyance from erroneous classifications.

2. Methodology

2.1. Support vector machine

The fundamental objective of SVM is to identify an optimal hyperplane (e.g., a line in 2D space or a plane in 3D space) that separates data points of different classes while maximizing the geometric margin between the hyperplane and the closest sample points (support vectors). This margin maximization enhances the model's generalization capability, reducing overfitting.

Key characteristics of SVM include:

(1) Support vector dependency

The final hyperplane is exclusively determined by support vectors (the critical samples nearest to the boundary). Non-support vectors (distant samples) have no impact, making SVM computationally efficient—especially for small-to-medium datasets.

(2) Kernel trick for nonlinear separation

For linearly inseparable data, SVM employs kernel functions (e.g., polynomial, Gaussian RBF) to implicitly map input features into a higher-dimensional space. This transformation converts complex nonlinear boundaries into linearly separable problems in the transformed space.

2.2. Dlib method

The design of the loss function significantly impacts training quality, particularly when the training dataset is limited. To address this, our study incorporates geometrical regularization and mitigates data imbalance by proposing the following loss function:

$$\mathrm{Loss} = \frac{1}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}\Upsilon_n||d_n^m||$$

where || || denotes a certain metric to measure the distance/error of the n-th landmark of the m-th input [15]. N is the number of landmarks to detect on each face. M is the number of images trained in each process [16].

Using a lightweight 68-point facial landmark model, we extracted seven key facial features: Jaw, mouth, nose, left eye and right eye, left eyebrow and right eyebrow.

These cropped features were then analyzed using multidimensional statistical methods to distinguish anger from a neutral baseline expression.

3. Results analysis

The integrated pattern recognition and convolutional neural network model achieved 86.2% overall accuracy in facial expression recognition. To intentionally reduce false positive identifications, we implemented a higher classification threshold of 0.8 for road anger detection. As defined by Hjørland, the recall rate represents the proportion of relevant cases successfully identified [16]. The complete cross-accuracy results for the pattern recognition process are presented in **Table 1**.

Table 1. Cross-accuracy for pattern recognition

	Classified to abnormal behavior	Classified to abnormal behavior	Recall rate
True abnormal behavior	2718	852	76.13%
True baseline	649	2921	81.82%
Accuracy rate	78.68%		

Convolutional Neural Networks (CNN) demonstrate significantly greater model transferability compared to traditional pattern recognition methods. Conventional pattern recognition systems often fail to accurately distinguish anger from baseline expressions due to their reliance on fixed-dimensional classification criteria. While increasing feature dimensions might seem beneficial, this approach risks overfitting due to the curse of dimensionality. Deep learning architectures address this limitation through multi-layered neural networks that effectively manage high-dimensional data.

The integrated Dlib-SVM process system achieved performance metrics of:

True positive rate: 76.13% (2718/3570 abnormal behavior cases correctly identified)

True negative rate: 81.82% (2921/3570 baseline cases correctly classified)

Overall accuracy: 78.68% (calculated as diagonal sum/total samples)

Despite decades of statistical research advancing emotion detection accuracy, current deep learning models, while demonstrating improved performance, still encounter challenging cases. Certain expressions remain ambiguous even to human evaluators, mirroring the model's classification difficulties. These persistent edge cases highlight the inherent complexity of accurate emotion recognition, particularly for subtle affective states like road anger.

4. Discussion

Road anger manifestations differ significantly from ordinary anger expressions. Unlike typical anger displays that are more pronounced and readily identifiable, road anger presents with subtler facial cues. This subdued nature necessitates more refined analytical techniques to differentiate road anger from neutral driving expressions. However, relying exclusively on camera footage for road anger detection proves challenging. A holistic approach incorporating multiple parameters becomes essential. Future research will integrate vehicle performance metrics like speed and acceleration, as these operational patterns effectively reflect aggressive driving behaviors. Additionally, wearable technology could monitor physiological indicators including heart rate, blood pressure, and blood oxygen levels. The synergistic combination of vehicle telemetry, biometric monitoring, and visual data will enable comprehensive road rage assessment.

While the current study focuses on data collection, it does not address the classification of varying anger intensity levels. This aspect, along with integrated analysis of vehicle parameters and driver biometrics, will constitute the next phase of research in traffic safety and road rage detection systems.

5. Conclusion

Aggressive driving behaviors stem from road rage. Detecting driver anger solely through camera input presents challenges due to individual variations in anger expression. Moreover, road-related anger displays tend to be more subtle compared to conventional anger manifestations. While typical anger expressions are often more pronounced, driving-induced rage appears less noticeable. Alternative detection methods incorporating heart rate monitoring and voice analysis can supplement visual cues, but reliance on additional sensors compromises practical implementation. This study introduces an effective computer vision approach combining dlib and SVM to identify driver anger with 86.2% accuracy. Although this laboratory-tested accuracy requires further field validation and data refinement, the integrated system successfully recognizes facial features across diverse challenging conditions, including noise, intense lighting, and backlighting. The dlib-SVM model demonstrates superior performance compared to existing anger detection systems.

The solution operates efficiently using only an embedded camera chip and local computing hardware without requiring supplementary processors. The standalone system communicates directly with vehicle control units, enabling real-time warnings or autonomous intervention in equipped vehicles without needing internet connectivity.

Key advantages of this approach include:

The camera-exclusive detection method enhances practical applicability across various scenarios.

Targeted facial feature extraction (eyes, eyebrows, mouth) minimizes biases from natural facial variations that might otherwise lead to false anger classifications.

The combined dlib-SVM framework, utilizing 68 facial landmarks for feature localization, ensures robust performance in diverse environmental conditions through comprehensive facial analysis and classification.

Disclosure statement

The author declares no conflict of interest.

References

[1] Qu W, Ge Y, Jiang C, et al., 2014, The Dula Dangerous Driving Index in China: An Investigation of Reliability and Validity. Accident Analysis Prevention, 64: 62–68.

[2] Underwood G, Chapman P, Wright S, et al., 1999, Anger While Driving. Transportation Research Part F Traffic Psychology & Behaviour, 2(1): 55–68.

[3] Gao H, Yuce A, Thiran J, 2014, Detecting Emotional Stress from Facial Expressions for Driving Safety, IEEE International Conference on Image Processing (ICIP), Paris, France, 5961–5965.

[4] Wan P, Wu C, Lin Y, et al., 2019, Driving Anger States Detection Based on Incremental Association Markov Blanket and Least Square Support Vector Machine. Discrete Dynamics in Nature and Society, 7: 1–17.

[5] Huang Y, Yang J, Liu S, et al., 2019, Combining Facial Expressions and Electroencephalography to Enhance Emotion Recognition. Future Internet, 11: 105–117.

[6] Zhang L, et al., 2021, EEG-Based Fatigue Detection During Simulated Driving: A Spectral Entropy Analysis. Transportation Research Part F, 78: 156–167.

[7] European Transport Safety Council, 2022, Regulating Fatigue in Professional Drivers: A Policy Review, ETSC, Brussels.

[8] World Health Organization, 2023, Global Status Report on Road Safety 2023, WHO Press, Geneva.

[9] Ministry of Transport, 2023, Annual Report on China's Motor Vehicle Ownership, People's Communications Press, Beijing.

[10] 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).

[11] Wang C, Dai Y, Zhou W, et al., 2020, A Vision-Based Video Crash Detection Framework for Mixed Traffic Flow Environment Considering Low-Visibility Condition. Journal of Advanced Transportation, 9194028.

[12] Chen X, Wang Y, 2023, Evaluation of Commercial Fatigue Monitoring Devices: A Meta-Analysis. IEEE Transactions on Intelligent Transportation Systems, 24(5): 1–12.

[13] International Transport Forum, 2024, Impact Report: 2021–23, ITF, Paris, viewed May 15, 2025.

[14] Kumari J, Rajesh R, Kumar A, 2016, Fusion of Features for the Effective Facial Expression Recognition. International Conference on Communication and Signal Processing, April 6–8, 2016, India.

[15] Guo X, Li S, Yu J, et al., 2019, PFLD: A Practical Facial Landmark Detector. Computer Vision and Pattern Recognition. arXiv:1902.10859.

[16] Hjorland B, 2009, The Foundation of the Concept of Relevance. Journal of the American Society for Information Science and Technology. 61(2): 217–237.