

SafeRider: Intelligent Driver Behaviour Monitoring System

Prof. Anuja Garande
Dept. of AI and Data Science
Zeal College of Engineering
Pune, India

Dhanshri Bajare
Artificial Intelligence and Data
Science
Zeal College of Engineering
Pune, India

Vaishnavi Jadhav
Artificial Intelligence and Data
Science
Zeal College of Engineering
Pune, India

Aditi Marathe
Artificial Intelligence and Data Science
Zeal College of Engineering
Pune, India

Anushka Kale
Artificial Intelligence and Data Science
Zeal College of Engineering
Pune, India

Abstract - Driver fatigue, distraction, and abnormal health conditions are major contributors to road accidents, highlighting the need for intelligent monitoring systems that assess driver behavior comprehensively [1],[2]. This paper presents an AI-driven driver behaviour monitoring system that integrates visual cues with numerical vehicle and health data to predict the driver's state in real time. The proposed system uses simulated facial image data representing conditions such as open eyes, closed eyes, yawning, and normal behaviour, along with numerical inputs including vehicle parameters (steering angle, speed) and driver health indicators (heart rate, body temperature, and blood pressure). A Convolutional Neural Network (CNN) is employed to analyze image-based features, while a Random Forest classifier processes numerical data to evaluate driving and health conditions. The outputs of both models are combined using a fusion logic mechanism to generate a final driver state prediction. Experimental results obtained using simulated datasets demonstrate that the multimodal fusion approach improves reliability and reduces false predictions compared to single-model systems [8], [17], [19]. The proposed system provides an effective framework for intelligent driver monitoring and can contribute to enhanced road safety and advanced driver assistance systems.

Index Term - Driver monitoring, CNN, Random Forest, Road safety, Machine learning, Flask, Deep learning, Text to speech.

I. INTRODUCTION

Road safety has become a critical concern due to the increasing number of traffic accidents caused by human factors such as driver fatigue, distraction, and abnormal physical conditions. Despite advancements in vehicle safety technologies, driver-related errors continue to be a major cause of accidents worldwide. Monitoring the driver's state in real time is therefore essential to reduce accident risks and improve overall driving safety.

Traditional driver behaviour monitoring systems

typically rely on a single source of information, such as vehicle dynamics, facial expressions, or physiological signals [8]. Systems based solely on vehicle data analyze parameters like speed and steering patterns but fail to capture the driver's

physical and cognitive condition. Similarly, vision-based systems focus on facial cues such as eye closure or yawning but may be affected by lighting conditions and occlusions [9], [13], [15]. Health-based monitoring systems, when used independently, provide physiological insights but lack contextual information about driving behaviour. As a result, unimodal approaches often produce incomplete or unreliable assessments of driver state.

This research presents an AI-driven driver behaviour monitoring system that combines visual information with numerical vehicle and driver health data to predict the driver's state accurately. The proposed system processes facial images to detect visual indicators such as open eyes, closed eyes, and yawning, while numerical parameters including steering angle, speed, heart rate, body temperature, and blood pressure are analyzed to assess driving and physical conditions. A fusion logic mechanism integrates the outputs of both models to generate a final driver state prediction. By leveraging multimodal data fusion, the proposed approach aims to improve reliability, reduce false predictions, and contribute to enhanced road safety and advanced driver assistance systems [8], [17], [19].

II. PROBLEM STATEMENT

Road accidents are frequently caused by driver fatigue, distraction, and abnormal physical conditions, which reduce alertness and increase the likelihood of unsafe driving [1], [8]. Existing driver monitoring systems often rely on a single source of information, such as visual cues, vehicle dynamics, or health parameters, analyzed independently. This unimodal approach limits the accuracy and reliability of driver state assessment, as it fails to capture the complex interaction

between driver behavior, vehicle movement, and physiological condition.

Moreover, systems that focus only on image-based detection may be affected by environmental factors, while numerical-data-based approaches lack direct observation of driver actions. Therefore, there is a need for an intelligent and integrated driver behaviour monitoring system that combines visual information with vehicle and driver health data[17], [19]. Such a system should be capable of analyzing multiple data modalities simultaneously and fusing their outputs to predict the driver's state accurately. Addressing this problem can enable early detection of unsafe driving conditions and support the development of advanced driver assistance systems aimed at improving road safety.

III. LITERATURE SURVEY

Driver behaviour monitoring has been an active area of research due to its importance in improving road safety and reducing accident rates. Several approaches have been proposed in the literature, primarily focusing on vision-based, vehicle-based, health-based, and hybrid monitoring systems.

Vision-based driver monitoring systems use computer vision techniques to analyze facial expressions and head movements for detecting driver fatigue and distraction. Methods such as eye closure detection, yawning detection, and head pose estimation have been widely studied. Convolutional Neural Networks (CNNs) have demonstrated strong performance in extracting spatial features from facial images and identifying drowsiness-related patterns[13], [15]. However, vision-based systems often suffer from limitations related to lighting conditions, camera angle variations, and occlusions such as sunglasses, which can reduce detection accuracy.

Vehicle-based monitoring systems analyze driving patterns using parameters such as speed, steering angle, lane deviation, and braking behavior. These systems are commonly used in telematics and fleet management applications to assess driving habits. While vehicle data provide valuable insights into driving dynamics, they do not directly reflect the driver's physical or cognitive state. As a result, unsafe conditions such as fatigue or health-related issues may go undetected when relying solely on vehicle parameters[4], [5].

Health-based and biometric monitoring approaches focus on physiological signals such as heart rate, body temperature, blood pressure, and electrocardiogram (ECG) readings to identify fatigue, stress, or abnormal driver conditions. These systems can effectively capture the driver's physical state, but when used independently, they lack contextual information about driving behavior and visual cues. Additionally, practical deployment may be affected by sensor availability and user comfort[11], [16]

IV. PROPOSED SOLUTION

The proposed solution aims to address the limitations of unimodal driver monitoring systems by introducing

an AI-driven, multimodal framework for accurate driver behaviour assessment. The system integrates visual information obtained from driver facial images with numerical data derived from vehicle dynamics and driver health parameters to provide a comprehensive evaluation of the driver's state. The proposed system processes facial images to detect visual indicators such as open eyes, closed eyes, yawning, and normal behaviour using a Convolutional Neural Network (CNN). These visual cues provide direct insight into driver alertness and attentiveness. In parallel, numerical data including vehicle parameters such as steering angle and speed, along with driver health indicators such as heart rate, body temperature, and blood pressure, are analyzed using a Random Forest classifier to assess driving behaviour and physiological condition

The proposed system processes facial images to detect visual indicators such as open eyes, closed eyes, yawning, and normal behaviour using a Convolutional Neural Network (CNN). These visual cues provide direct insight into driver alertness and attentiveness. In parallel, numerical data including vehicle parameters such as steering angle and speed, along with driver health indicators such as heart rate, body temperature, and blood pressure, are analyzed using a Random Forest classifier to assess driving behaviour and physiological condition.

To enhance reliability, the outputs of both models are combined using a fusion logic mechanism. This fusion-based decision-making process evaluates the predictions from the image-based and numerical-data-based models to generate a final driver state classification. If both models indicate unsafe or abnormal behaviour, the system identifies a high-risk driver condition. If only one model detects abnormality, a warning-level state is generated, while normal behaviour is predicted when both models indicate safe conditions.

The proposed AI-driven driver behaviour monitoring system is designed using a multimodal architecture that integrates visual information with numerical vehicle and driver health data to achieve reliable driver state prediction. The overall system architecture, as illustrated in Fig. 1, consists of multiple interconnected modules responsible for data acquisition, preprocessing, model-based analysis, fusion logic, and final decision making.

V. SYSTEM ARCHITECTURE

A. Data Acquisition Module

The system acquires input data from two primary sources:
Visual Data Acquisition : Facial images of the driver are collected using a camera positioned inside the vehicle. The image dataset includes different driver states such as open eyes, closed eyes, yawning, and not yawning, which represent visual indicators of alertness and fatigue.

Numerical Data Acquisition : Numerical inputs include both vehicle-related parameters and driver health parameters. Vehicle data consist of features such as steering angle and speed, while driver health data include parameters such as heart rate, body temperature, and blood pressure. These parameters provide insights into the driver's driving behaviour and physical condition.

B. Preprocessing Module

To ensure effective model performance, the acquired data undergo preprocessing before being passed to the learning models.

Image Preprocessing : Facial images are resized to a fixed resolution of 224×224 pixels, normalized, and converted into numerical arrays suitable for CNN input.

Numerical Data Preprocessing : Vehicle and health data are cleaned to handle missing or inconsistent values. Feature scaling is applied to normalize the numerical parameters and maintain uniformity across input features.

C. Image-Based Behaviour Analysis

The preprocessed facial images are fed into a Convolutional Neural Network (CNN), which automatically extracts spatial features related to driver facial expressions. The CNN classifies the visual input into categories such as open eyes, closed eyes, yawning, or normal behaviour. This module focuses on identifying visual signs of drowsiness or inattentiveness.

D. Numerical Data Analysis

The preprocessed numerical data are processed using a Random Forest classifier. This model analyzes vehicle dynamics and driver health parameters to classify the driver's condition into behavioural states such as attentive, drowsy, or unsafe. The ensemble-based nature of the Random Forest model improves robustness and reduces sensitivity to noise in the data.

E. Fusion Logic and Decision Module

The fusion logic module integrates the outputs obtained from the CNN-based image analysis and the Random Forest-based numerical analysis. A rule-based fusion mechanism is employed to combine predictions from both models. If both models indicate unsafe or abnormal conditions, the system predicts a high-risk driver state. If only one model detects abnormality, a warning-level state is generated. When both models indicate normal behaviour, the driver state is classified as safe.

F. Driver State Prediction

Based on the fusion logic, the final driver state prediction is generated. This multimodal architecture improves detection accuracy and reliability compared to single-modality systems. The system architecture provides a scalable and efficient framework for real-time driver behaviour monitoring and supports the development of intelligent transportation systems and advanced driver assistance systems.

System Architecture of AI-Driven Driver Behaviour Monitoring System

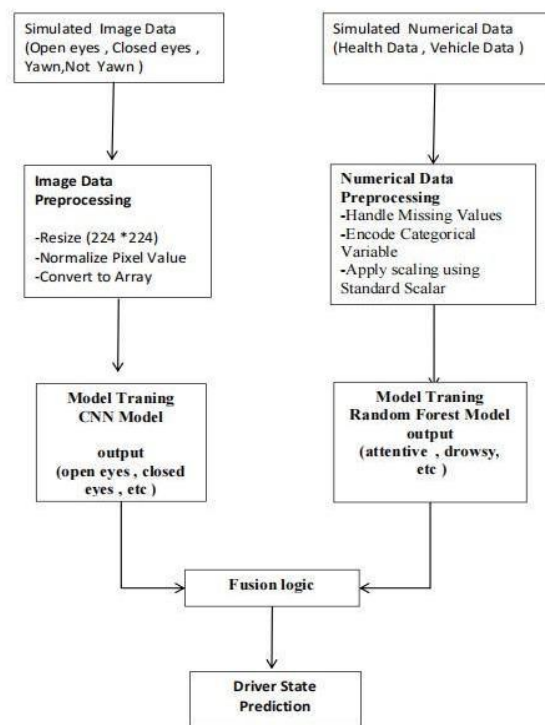


Fig. 1 System Architecture

VI. METHODOLOGY

The proposed methodology aims to develop an intelligent driver behaviour monitoring system by integrating visual and numerical data through a multimodal learning approach. The system combines deep learning and machine learning techniques[8] to analyze driver facial images, vehicle parameters, and driver health indicators. The overall methodology consists of five major stages: data collection, preprocessing, model development, fusion logic implementation, and driver state prediction.

A. Data Collection

Two categories of datasets are used in the proposed system:

Image Dataset:The image dataset consists of simulated facial images representing different driver states, including open eyes, closed eyes, yawning, and not yawning. These images are used to train and evaluate the CNN-based image classification model.

Numerical Dataset:The numerical dataset includes both vehicle-related and driver health-related parameters. Vehicle data consist of features such as steering angle and speed, while health data include parameters such as heart rate, body temperature, and blood pressure. These numerical values represent driving behaviour and physiological conditions of the driver.

B. Data Preprocessing

To ensure consistent and effective learning, preprocessing is applied to both image and numerical data.

Image Preprocessing:All facial images are resized to 224×224 pixels and normalized to scale pixel intensity values. These steps reduce computational complexity and improve the CNN's feature extraction capability.

Numerical Data Preprocessing:Numerical datasets are cleaned by removing or handling missing values. Feature scaling is applied to normalize the data and maintain equal contribution of each parameter during model training.

C. Model Development

Two independent models are developed to handle different data modalities:

CNN-Based Image Classification Model:A Convolutional Neural Network is designed to automatically extract spatial features from driver facial images. The CNN processes the preprocessed images and classifies them into visual driver states such as open eyes, closed eyes, yawning, or normal behaviour [13], [15].

Random Forest-Based Numerical Classification Model:A Random Forest classifier is trained using the numerical vehicle and health data. The ensemble-based model analyzes the input parameters and classifies the driver's condition into behavioural states such as attentive or unsafe[8].

D. Fusion Logic Implementation

The fusion logic module integrates the predictions obtained from the CNN and Random Forest models. A rule-based fusion approach is employed to combine the outputs of both models. If both models detect abnormal or unsafe behaviour, the system generates a high-risk driver state. If only one model indicates abnormality, a warning-level state is produced. When both models indicate normal conditions, the driver state is classified as safe[17], [19]. This fusion strategy enhances reliability and reduces false predictions.

E. Driver State Prediction and Output

The final driver state prediction is generated based on the fusion logic output. By combining visual cues with numerical vehicle and health parameters, the proposed methodology provides a comprehensive assessment of driver behaviour. The multimodal approach improves detection accuracy and robustness compared to single-modality systems, making it suitable for intelligent driver monitoring applications.

VII. IMPLEMENTATION

The proposed AI-Driven Driver Behavior Assessment System is implemented by combining deep learning and machine learning models in a single software framework. Python is used for backend development with Flask, while HTML,CSS,Javascript is used for the frontend to enable smooth communication and real-time result display.The overall implementation is divided into six main stages: data collection, data preprocessing, model training, system integration, frontend development, and testing.

A. Data Collection and Preparation

Since collecting real driver monitoring data is difficult due to privacy and safety issues, simulated datasets are used for experimentation. The datasets are divided into two types:

Visual Dataset :

This dataset consists of labeled facial images representing different driver states such as open eyes, closed eyes, yawning, and not yawning. These images are used to train and test the CNN model. All images are resized to 224×224 pixels and normalized to improve model performance.

Numerical Dataset :

This dataset contains numerical features such as vehicle speed, steering angle, heart rate, and stress level. The data is normalized and encoded before being fed into the Random Forest classifier. The output labels represent driver states such as attentive, drowsy, and distracted.

B. Model Training

Convolutional Neural Network (CNN) :

The CNN model is developed using the TensorFlow-Keras framework. It consists of convolutional layers, pooling layers, dropout layers, and fully connected layers to extract facial features effectively. The model is trained using the Adam optimizer with categorical cross-entropy loss. CNN achieved a training accuracy of more than 94% on the image dataset[22].

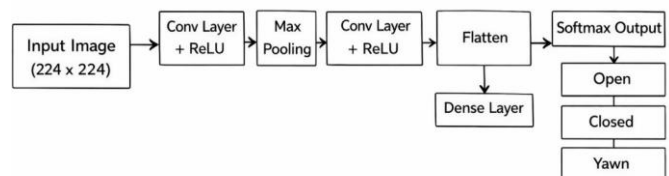


Fig . 2 CNN Architecture for Facial Behaviour Detection

Random Forest Classifier :

The Random Forest model is implemented using the Scikit-learn library. It uses multiple decision trees to classify driver behavior based on numerical inputs. The final prediction is obtained using majority voting. This model achieved an average accuracy of 91% on validation data[23].

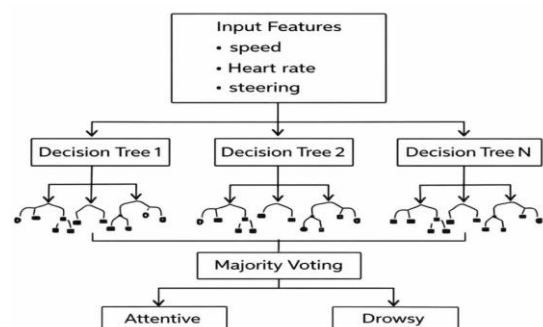


Fig . 3 Random Forest Classifier for Numerical Data

Both trained models are saved for later use — driver_custom_cnn.h5 for the CNN model and random_forest_driver.pkl for the Random Forest model.

C. Backend Integration (Flask API)

The Flask backend serves as the main processing unit of the system. It loads both trained models during initialization and handles all prediction requests. The backend provides three main API endpoints:

/predict_image – Facial images prediction using the CNN

/predict_numerical – Numerical Prediction using the Random Forest model

/predict_fusion – Gives final prediction based on image data and numerical data

The backend ensures fast predictions and sends results to the frontend in JSON format.

D. Frontend Development

The frontend provides a simple and interactive user interface. Users can upload driver images and enter numerical values such as speed and heart rate through input forms. The frontend sends this data to the Flask API using HTTP POST requests. The prediction results and alert messages such as Normal, Warning, or Critical Alert are displayed instantly. Visual indicators like progress bars and color-coded alerts (green, yellow, and red) improve clarity and user understanding.

E. Fusion Logic Implementation

Fusion logic is used to combine predictions from both models to improve accuracy. If the CNN detects closed eyes or yawning and the Random Forest predicts drowsy or distracted behavior, a Critical Alert is generated.

If only one model indicates risky behavior, a Warning is shown. If both models indicate normal behavior, the system outputs Driver Alert (Normal).

This approach reduces false alarms and improves overall reliability.

F. Testing and Validation

The complete system is tested using simulated and unseen data. Backend responses are verified using Postman, and real-time results are checked through the User Interface. The integrated system achieved an overall accuracy of 96%, confirming effective driver state detection and real-time alert generation.

VIII. DESIGN AND RESULT

A. CNN Model Results

The Convolutional Neural Network (CNN) is used to classify driver facial states such as open eyes, closed eyes, yawning, and not yawning. The performance of the CNN model is evaluated using a confusion matrix and standard classification metrics. The model achieves an overall accuracy of 95%, indicating effective detection of facial features related to driver alertness.

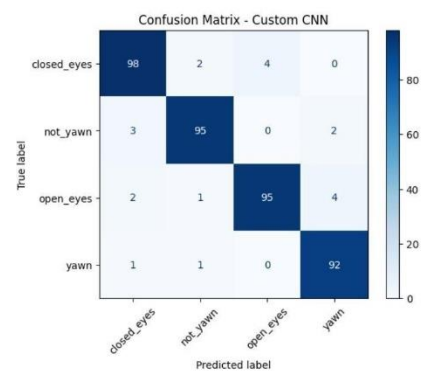


Fig. 4 Confusion Matrix of CNN Model

B. Random Forest Model Results

The Random Forest model is employed to analyze sequential numerical data related to vehicle dynamics and driver health parameters. The model classifies driving behaviour into categories such as attentive, distracted, drowsy, aggressive, and eco-driving. The RF model achieves an overall accuracy of approximately 98%, reflecting strong performance in capturing temporal patterns in the data.

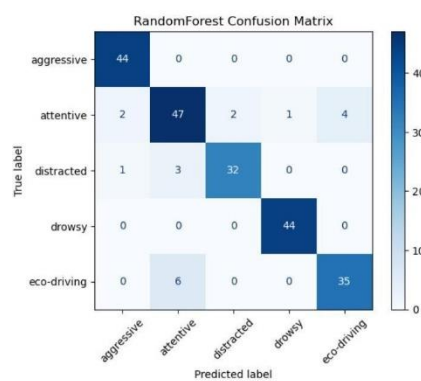


Fig. 5 Confusion Matrix for Random Forest

IX. CONCLUSION

This work presents an AI-based hybrid system for real-time driver behavior monitoring using facial image analysis and numerical vehicle and health data. The combination of CNN for visual features and Random Forest for numerical data allows accurate detection of driver states such as attentive, drowsy, and distracted.

The fusion-based decision approach improves accuracy [17],[19] and reduces false detections compared to single-model systems. The Flask backend efficiently manages predictions, while the frontend provides a clear and user-friendly interface. Experimental results using simulated data achieved an overall accuracy of 96%, demonstrating the system's reliability and effectiveness for driver safety applications.

X. FUTURE SCOPE

Although the current system uses simulated data, several enhancements can be made for real-world deployment :

1) Integration with Real-Time Sensors :

Use IoT devices such as smartwatches, ECG sensors, and vehicle telematics to collect live driver data.

2) Advanced Deep Learning Models :

Implement transfer learning models like ResNet or EfficientNet to improve facial analysis under different lighting and environmental conditions.

3) Edge and Cloud Deployment :

Deploy the system on edge devices such as Raspberry Pi or Jetson Nano for in-vehicle processing, with cloud support for large-scale data analysis.

4) Adaptive Learning :

Enable the system to learn continuously and adapt to different drivers and driving behaviors.

5) Integration with ADAS :

Connect the system with Advanced Driver Assistance Systems to automatically trigger safety actions such as speed control or alert mechanisms[6], [7].

XI. REFERENCES

1. A. G. Wheaton, "Drowsy driving and risk behaviors—10 states and Puerto Rico 2011–2012," *MMWR Morb Mortal Wkly Rep*, vol. 63, no. 26, pp. 557–562, 2014.
2. *NHTSA*. Accessed: Feb. 27, 2024. [Online]. Available: <https://www.nhtsa.gov/risky-driving/drowsy-driving>
3. *Canalys*. Accessed: Feb. 20, 2024. [Online]. Available: <https://www.canalys.com/newsroom/canalys-autonomous-driving-starts-to-hit-mainstream-as-35-million-new-cars-had-level-2-features-in-q4-2020>
4. O. Carsten, F. C. H. Lai, Y. Barnard, A. H. Jamson, and N. Merat, "Control task substitution in semiautomated driving: Does it matter what aspects are automated?" *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 54, no. 5, pp. 747–761, Oct. 2012, doi: 10.1177/0018720812460246.
5. J. W. Jenness, N. D. Lerner, S. Mazor, J. Osberg, and S. Tefft, "Use of advanced in-vehicle technology by young and older early adopters. Survey results on adaptive cruise control systems," NHTSA, U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. DOT HS 810 917, 2008.
6. *IDTechEx*. Accessed: Feb. 15, 2024. [Online]. Available: <https://www.idtechex.com/ja/research-article/regulations-drivers-for-mandating-driver-monitoring-systems/30322>
7. *U.S. Congress*. Accessed: Feb. 15, 2024. [Online]. Available: <https://www.congress.gov/bill/117th-congress/senate-bill/1406/text>
8. M. Q. Khan and S. Lee, "A comprehensive survey of driving monitoring and assistance systems," *Sensors*, vol. 19, no. 11, p. 2574, Jun. 2019, doi: 10.3390/s19112574.
9. S. Titare, S. Chinchghare, and K. N. Hande, "Driver drowsiness detection and alert system," *Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol.*, vol. 7, pp. 583–588, Jun. 2021, doi: 10.32628/cseit2173171.
10. C. Solomon and Z. Wang, "Driver attention and behavior detection with Kinect," *J. Image Graph.*, vol. 3, no. 2, pp. 1–6, 2015. [Online]. Available: <https://www.joig.net/index.php?m=content&c=index&a=show&catid=42&id=106>
11. J. D. Ortega, N. Köse, P. N. Cañas, M.-A. Chao, A. Unnervik, M. Nieto, O. Otaegui, and L. Salgado, "DMD: A large-scale multi-modal driver monitoring dataset for attention and alertness analysis," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Jan. 2020, pp. 387–405, doi: 10.1007/978-3-030-66823-5_23.
12. B. K. Savas and Y. Becerikli, "Real time driver fatigue detection system based on multi-task ConNN," *IEEE Access*, vol. 8, pp. 12491–12498, 2020, doi: 10.1109/ACCESS.2020.2963960.
13. E. Vural, "Drowsy driver detection through facial movement analysis," in *Proc. Int. Workshop Hum.-Comput. Interact.*, vol. 4796, Cham, Switzerland: Springer, pp. 6–18, doi: 10.1007/978-3-540-75773-3_2.
14. K. Dwivedi, K. Biswaranjan, and A. Sethi, "Drowsy driver detection using representation learning," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Feb. 2014, pp. 995–999, doi: 10.1109/IADCC.2014.6779459.
15. W. Zhang, Y. L. Murphey, T. Wang, and Q. Xu, "Driver yawning detection based on deep convolutional neural learning and robust nose tracking," in *Proc. IJCNN*, Jul. 2015, pp. 1–8, doi: 10.1109/ijcnn.2015.7280566.
16. Y. Ji, S. Wang, Y. Zhao, J. Wei, and Y. Lu, "Fatigue state detection based on multi-index fusion and state recognition network," *IEEE Access*, vol. 7, pp. 64136–64147, 2019, doi: 10.1109/ACCESS.2019.2917382.
17. K. Zhang, S. Wang, N. Jia, L. Zhao, C. Han, and L. Li, "Integrating visual large language model and reasoning chain for driver behavior analysis and risk assessment," *Accident Anal. Prevention*, vol. 198, Apr. 2024, Art. no. 107497, doi: 10.1016/j.aap.2024.107497.
18. J. Chen, S. Dey, L. Wang, N. Bi, and P. Liu, "Attention-based multi-modal multi-view fusion approach for driver facial expression recognition," *IEEE Access*, vol. 12, pp. 137203–137221, 2024, doi: 10.1109/ACCESS.2024.3462352.
19. J.-W. Kim, C. I. Nwakanma, D.-S. Kim, and J.-M. Lee, "Intelligent face recognition on the edge computing using neuromorphic technology," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2021, pp. 514–516, doi: 10.1109/ICOIN50884.2021.9333967.
20. S. Y. Nikouei, Y. Chen, S. Song, R. Xu, B.-Y. Choi, and T. Faughnan, "Smart surveillance as an edge network service: From harr-cascade, SVM to a lightweight CNN," in *Proc. IEEE 4th Int. Conf. Collaboration Internet Comput. (CIC)*, Oct. 2018, pp. 256–265, doi: 10.1109/CIC.2018.00042.
21. <https://www.geeksforgoeks.org/machine-learning/introduction-convolutional-neural-network/>
22. <https://www.geeksforgoeks.org/machine-learning/random-forest-algorithm-in-machine-learning/>