International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

Drowsiness Detection Using Convolutional Neural Network

Mr. Nagesh H B *Electronics and Commn. Engg. ACS College of Engineering*
Bangalore, India
nagesh.murthy@gmail.com

Likith Gowda K N
Dept. of Electronics and Commn. Engg.
ACS College of Engineering
Bangalore, India
likithgowdakn5032@gmail.com

Pretham L
Dept. of Electronics and Commn. Engg.
ACS College of Engineering Bangalore, India
prethaml2003@gmail.com

Sharath S
Dept. of Electronics and Commn. Engg.
ACS College of Engineering
Bangalore, India
sharathsgowda15@gmail.com

Sujay Singh M Rajputh
Dept. of Electronics and Commn. Engg ACS College of Engineering Bangalore, India
sujaysinghmrajputh@gmail.com

Abstract— Driver drowsiness is one of the major causes of road accidents, especially during long-distance and night-time driving. Continuous monitoring of driver alertness using non- intrusive techniques has therefore become essential for improving road safety. This paper presents a real-time drowsiness detection system based on facial feature analysis using a Convolutional Neural Network (CNN). The proposed system captures live video through a standard webcam and processes facial features such as eye closure and yawning to determine the driver's alertness level. Facial landmarks are extracted to compute Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), which serve as key indicators of fatigue. A multi-scale CNN architecture is employed to learn both fine and coarse facial patterns, improving detection accuracy under varying lighting conditions and face orientations. When drowsiness is detected, the system immediately generates visual and audio alerts to warn the driver. The experimental results demonstrate that the proposed approach operates effectively in real time using low-cost hardware, making it suitable for practical deployment in intelligent transportation and safety systems.

Keywords— Drowsiness Detection, Convolutional Neural Network (CNN), Facial Feature Analysis, Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), Real-Time Monitoring, Driver Safety.

I. INTRODUCTION

Road safety has become a critical global concern due to the increasing number of accidents caused by human factors, among which driver fatigue and drowsiness play a significant role. Long driving hours, inadequate sleep, night-time travel, and monotonous road conditions often reduce a driver's level of alertness, leading to delayed reactions and impaired decision-making. According to various traffic safety studies, drowsiness-related accidents are more severe than other types of accidents because drivers fail to take timely corrective actions. Therefore, continuous monitoring of driver alertness has become an essential component of modern transportation safety systems.

Conventional methods for detecting driver drowsiness include physiological signal monitoring such as electroencephalogram (EEG), electrocardiogram (ECG), and heart rate analysis. Although these methods provide accurate measurements, they require wearable sensors that are uncomfortable and impractical for real-world driving environments. Vehicle-based approaches that analyze steering behavior, lane deviation, and pedal usage have also been explored; however, these methods are highly dependent on road conditions, vehicle type, and driving style. As a result, there is a growing demand for a non-intrusive, reliable, and cost-effective solution for detecting driver drowsiness.

Vision-based approaches have gained significant attention due to their non-contact nature and ability to directly analyze facial behavior. Facial cues such as eye closure duration, blink frequency, head pose, and yawning patterns serve as strong indicators of fatigue. Advances in computer vision and deep learning have enabled accurate extraction and analysis of these facial features using camera-based systems. Among various deep learning models, Convolutional Neural Networks (CNNs) have demonstrated superior performance in image-based feature extraction and classification tasks, making them suitable for real-time drowsiness detection.

Recent research has shown that combining traditional facial feature metrics with deep learning techniques improves detection accuracy and robustness. Metrics such as Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), derived from facial landmark detection, provide reliable measures of eye closure and yawning behavior. However, single-scale CNN models often struggle to generalize under varying lighting conditions, facial orientations, and user demographics. To address these challenges, multi-scale CNN architectures have been introduced to capture both fine-grained and global facial features, thereby enhancing model performance.

In this paper, a real-time drowsiness detection system using a Convolutional Neural Network with a multi-scale architecture is proposed. The system captures live video using a standard webcam and performs facial detection and landmark extraction to analyze eye and mouth movements. EAR and MAR values are continuously monitored and processed by the CNN model to determine the driver's alertness level. When drowsiness is detected, the system immediately triggers visual and auditory alerts to warn the driver and prevent potential accidents. The proposed system operates efficiently on low-cost hardware, making it suitable for practical deployment in intelligent transportation systems and driver assistance applications.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

II. Literature Survey

Driver drowsiness detection has been an active area of research for several decades due to its importance in improving road safety. Existing approaches for detecting driver fatigue can be broadly categorized into physiological-based, vehicle-based, and vision-based methods. Each category offers unique advantages but also presents certain limitations when applied in real-world driving environments.

Physiological-based approaches monitor biological signals such as electroencephalogram (EEG), electrooculogram (EOG), electrocardiogram (ECG), and heart rate variability to assess a driver's level of alertness. These methods provide high accuracy since physiological signals directly reflect fatigue levels. However, they require the use of wearable sensors or electrodes, which can cause discomfort and distraction to the driver. Additionally, such systems are expensive and impractical for continuous real-time monitoring in everyday driving conditions.

Vehicle-based methods analyze driving behavior and vehicle dynamics, including steering wheel movements, lane departure patterns, acceleration, and braking behavior. While these approaches do not require direct interaction with the driver, their performance is heavily influenced by external factors such as road conditions, vehicle type, weather, and individual driving styles. As a result, these methods often suffer from reduced reliability and limited generalizability across different driving environments.

Vision-based approaches have emerged as a promising alternative due to their non-intrusive nature and direct analysis of driver behavior. These systems rely on camera-based monitoring to observe facial cues such as eye blinking, eye closure duration, head pose, and yawning frequency. Early vision-based systems utilized traditional image processing techniques, including Haar cascade classifiers and threshold-based methods, to detect facial features. However, these techniques struggled under variations in lighting conditions, facial orientation, and occlusions.

With advancements in deep learning, Convolutional Neural Networks (CNNs) have been widely adopted for facial analysis tasks due to their ability to automatically learn hierarchical features from images. Several studies demonstrated that CNN-based models significantly outperform traditional methods in detecting eye states and facial expressions related to drowsiness. Pre-trained models and transfer learning techniques further improved detection accuracy while reducing training time.

Recent research has focused on combining facial landmark-based features with deep learning models to enhance robustness. Metrics such as Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) have been used to quantify eye closure and yawning behavior effectively. Multi-scale CNN architectures have also been proposed to capture both local and global facial features, improving performance under diverse environmental conditions. However, some existing systems require high computational resources or lack real-time capability.

The proposed system builds upon these prior works by integrating facial landmark-based feature extraction with a multi-scale CNN model to achieve accurate and real-time drowsiness detection using low-cost hardware. This approach addresses key limitations of existing methods and offers a practical solution for real-world driver monitoring applications.

III. Proposed System

The proposed system aims to detect driver drowsiness in real time by analyzing facial features using a vision-based approach combined with deep learning techniques. The system is designed to be non- intrusive, cost-effective, and suitable for real-world deployment in intelligent transportation and driver monitoring systems. By continuously monitoring facial cues such as eye closure and yawning, the system identifies early signs of fatigue and provides timely alerts to prevent accidents.

The overall architecture of the proposed system is illustrated in **Fig. 1**, which presents the major functional components and their interactions. The system primarily consists of a video acquisition module, face detection module, facial landmark extraction module, feature computation module, CNN-based classification module, and alert generation module. Each component works sequentially to ensure smooth and continuous monitoring of the driver's alertness state.

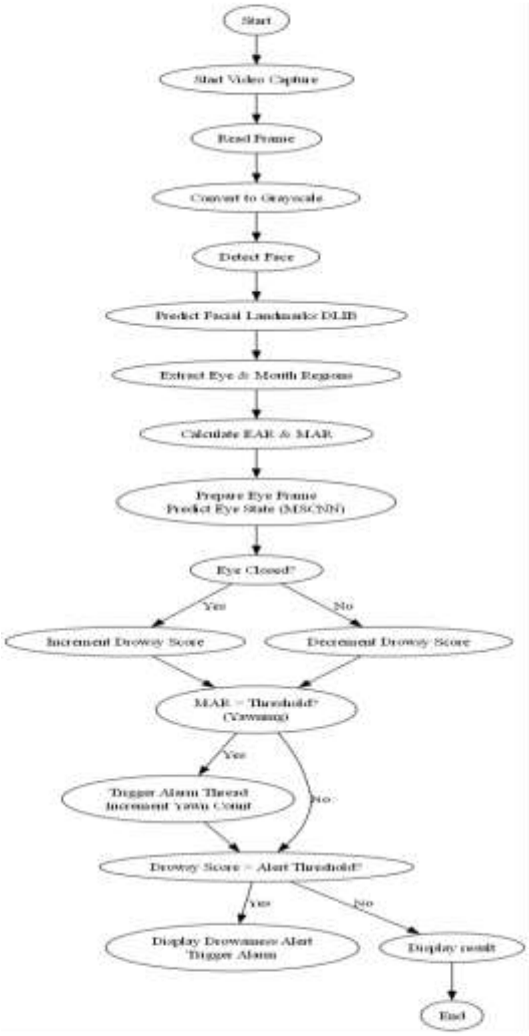

Fig 1: Flowchart Of Methodology

The system begins with real-time video acquisition using a standard USB webcam positioned to capture the driver's face. The webcam continuously records video frames at a fixed frame rate. Each frame is extracted from the video stream and passed to the preprocessing stage. Preprocessing includes resizing the frame to a suitable resolution and converting it into grayscale to reduce computational complexity and improve processing speed without losing essential facial details.

After preprocessing, face detection is performed on each frame to locate the driver's face. This step ensures that further processing is applied only when a face is present in the frame. If no face is detected, the system ignores the frame and continues capturing subsequent frames. Once a face is detected, the region of interest corresponding to the face is extracted for further analysis.

Facial landmark detection is then applied to the detected face region. This step identifies key facial points around important regions such as the eyes and mouth. These landmark points provide a precise geometric representation of facial movements and expressions. The facial landmark detection process highlighting the points used for eye and mouth analysis. Accurate landmark detection is crucial, as the reliability of drowsiness detection depends heavily on the precision of these points.

Using the extracted landmarks, fatigue-related features are computed. The Eye Aspect Ratio (EAR) is calculated to measure the openness of the eyes. The EAR provides a robust measure of eye closure by comparing vertical and horizontal distances between eye landmarks. A sustained decrease in EAR over multiple consecutive frames indicates prolonged eye closure, which is a strong indicator of drowsiness. To improve reliability, the EAR is computed separately for both eyes, and the average value is used for decision-making.

The Eye Aspect Ratio (EAR) is used to measure eye openness and is calculated using six eye landmark points, as commonly defined in facial landmark models.

Let P1, P2, P3, P4, P5, and P6 represent the eye landmark coordinates. The EAR is defined as:

$$EAR = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2\ X\ ||P_1 - P_4||}$$

where || . || denotes the Euclidean distance between two points. The EAR is comp final EAR value is obtained by averaging both values:

$$EAR_{AVG} = \frac{EAR_{left} + EAR_{right}}{2}$$

A continuous drop in the EAR value below a predefined threshold indicates prolonged

In addition to eye closure detection, yawning behavior is analyzed using the Mouth Aspect Ratio (MAR). The MAR is calculated using selected mouth landmarks to measure the extent of mouth opening.

A significant increase in MAR over time indicates yawning, which is commonly associated with fatigue and reduced alertness. Continuous monitoring of MAR values enables the system to detect repeated or prolonged yawning events.

The Mouth Aspect Ratio (MAR) is used to detect yawning behavior based on mouth landmark points. Let P1, P2, ……P8 denote selected landmark points around the mouth region.

The MAR is calculated as:

$$MAR = \frac{||P_3 - P_7|| + ||P_4 - P_6|| + ||P_5 - P_8||}{3\ X\ ||P_1 - P_2||}$$

An increase in MAR value over consecutive frames indicates significant mouth opening, which is associated with yawning and driver fatigue.

The computed EAR and MAR values serve as key inputs to the classification stage. These features, along with facial information extracted from the video frames, are processed using a Convolutional Neural Network (CNN). The CNN model is designed to automatically learn spatial features from facial images and distinguish between alert and drowsy states. A multi-scale CNN architecture is employed to capture both fine-grained and global facial features, improving robustness under varying lighting conditions, facial orientations, and individual differences.

The CNN classifier analyzes the input features and assigns the driver's state as either alert or drowsy. The classification decision is based on learned patterns obtained during the training phase using labeled datasets containing drowsy and non-drowsy facial images. The integration of handcrafted features such as EAR and MAR with deep learning-based feature extraction enhances detection accuracy and reduces false positives.

The operational flow of the proposed system is illustrated in **Fig. 2**, which depicts the step-by-step execution from video capture to alert generation. The system continuously processes incoming frames in real time, ensuring minimal latency between drowsiness detection and alert activation. This real-time capability is essential for safety- critical applications such as driver monitoring.

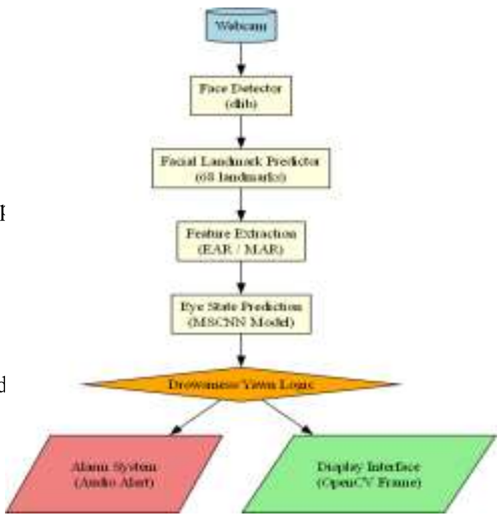

Fig 2 : Flowchart Of Proposed System

When the system detects drowsiness beyond predefined thresholds, it immediately triggers alert mechanisms. These alerts include visual warnings displayed on the screen and audible alarms to regain the driver's attention. The alert generation output, including warning messages and notifications alerts are sent. These alerts are designed to be simple yet effective, ensuring that the driver is promptly warned of fatigue.

Overall, the proposed system provides an effective, non-intrusive, and scalable solution for real-time driver drowsiness detection. By combining facial landmark-based feature extraction with CNN-based classification, the system achieves reliable performance using low-cost hardware. The modular design allows easy integration with Advanced Driver Assistance Systems (ADAS) and provides scope for future enhancements such as temporal modeling, embedded deployment, and night-time detection.

IV. Hardware and Software Requirements

The proposed drowsiness detection system is designed to operate using standard and easily available hardware components, ensuring cost effectiveness and ease of deployment. A laptop or personal computer is required as the primary development and execution platform for coding, training, and testing the Convolutional Neural Network model. The system requires a minimum Intel Core i5 processor, while an Intel Core i7 processor is recommended to achieve better real-time performance during continuous video processing and model inference. A minimum of 8 GB RAM is necessary to ensure smooth execution of Python programs, OpenCV-based video processing, and deep learning frameworks without performance lag.

For storage, a minimum of 512 GB SSD is recommended to allow faster boot times, quicker access to datasets, trained models, libraries, and development tools. A standard USB webcam is used for capturing live facial video of the driver in real time. This camera-based setup eliminates the need for intrusive or expensive sensors, making the system practical for real-world applications. Additionally, a stable internet connection is required for downloading software libraries, accessing documentation, updating packages, and using version control platforms such as GitHub.The software components used in the proposed system are listed below:

- **Python:** Used as the primary programming language for implementing computer vision, feature extraction, and CNN- based classification.
- **Machine Learning Libraries:** TensorFlow and Keras are used for designing, training, and deploying the Convolutional Neural Network model.
- **OpenCV and dlib:** OpenCV handles real-time video capture and image preprocessing, while dlib is used for accurate facial landmark detection.
- **Real-Time Processing Support:** Threading is optionally used to improve processing speed and responsiveness.
- **Alert Libraries:** Playsound or Pygame is used to generate audio alerts when drowsiness is detected.
- **Anaconda PowerShell Prompt:** Used for environment setup, package management, and dependency control.

V. Results and discussion

The proposed Facial Feature–Based Drowsiness Detection System was tested extensively under real-time conditions to evaluate its effectiveness in identifying driver fatigue. The evaluation was carried out using a standard USB webcam and mid-range computing hardware. The system was tested with multiple subjects to analyze its performance across different facial structures, eye shapes, and expressions. The primary focus of the evaluation was on eye closure detection, yawning detection, alert generation, and real-time responsiveness.

During the testing phase, the system continuously captured live video frames and successfully detected the driver's face in most scenarios. Facial landmarks corresponding to the eyes and mouth were accurately extracted in real time, allowing reliable computation of the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR). These features were then passed to the CNN classifier to determine the driver's alertness state.

Normal Alert Condition

Under normal driving conditions, when the driver was fully alert, the system maintained stable EAR and MAR values. The eyes remained open, and no excessive mouth opening was observed. The CNN classifier correctly classified these frames as non-drowsy. **Fig 3** shows the system output during an alert condition, where the live video feed displays normal facial behavior without any warning or alarm messages. This result confirms that the system does not generate false alerts during normal operation.

Fig 3 : Detection of Eye Open

Drowsiness Detection Performance

When the driver began showing signs of fatigue such as prolonged eye closure or repeated yawning, noticeable changes were observed in the computed EAR and MAR values. The EAR dropped below the predefined threshold for consecutive frames, indicating eye closure, while the MAR increased significantly during yawning events. The CNN classifier effectively detected these patterns and classified the driver's state as drowsy. **Fig 4** illustrates the system output during drowsiness detection, where a warning message is displayed on the screen indicating reduced alertness.

This result demonstrates the capability of the system to accurately identify early symptoms of fatigue before the driver completely loses alertness. The integration of facial landmark-based features with CNN-based classification significantly improves detection reliability and reduces misclassification.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

Fig 4 : Detection of Eye Closure

Alert and Warning Mechanism

Once drowsiness was detected continuously over a predefined number of frames, the system triggered an alert mechanism to warn the driver. This included both visual warnings and an audible alarm to regain the driver's attention. **Fig 5** shows the alert generated by the system, highlighting the warning message and alarm activation. The alerts were generated with minimal delay, demonstrating the real-time capability of the system.

The alert mechanism plays a crucial role in preventing fatigue-related accidents by providing immediate feedback to the driver. The results confirm that the system responds promptly to drowsiness detection and effectively alerts the driver without significant latency.

Fig 5 : Alert Message through E-mail

Overall Performance Analysis

Overall, the system demonstrated reliable performance in real-time drowsiness detection using low-cost hardware. The combination of EAR, MAR, and CNN-based classification allowed accurate detection of both eye closure and yawning behaviors. The system performed well under normal lighting conditions and moderate head movements. However, minor performance degradation was observed under poor lighting or partial facial occlusion, indicating potential areas for future improvement.

VI. Conclusion

This project presents a real-time, non-intrusive drowsiness detection system based on facial feature analysis and Convolutional Neural Networks. By monitoring eye closure and yawning behavior through Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), the system effectively identifies early signs of driver fatigue.

The integration of facial landmark–based feature extraction with CNN-based classification enables accurate differentiation between alert and drowsy states. Experimental results demonstrate that the system operates reliably using a standard webcam and mid-range hardware, providing timely visual and audio alerts to warn the driver. The cost-effective design and real-time performance make the proposed approach suitable for practical deployment in intelligent transportation and driver monitoring applications.

While the current system demonstrates promising performance, several enhancements can be explored in future work to further improve robustness and applicability. Performance under low-light and night-time conditions can be enhanced by incorporating infrared-based imaging techniques. Temporal modeling approaches such as LSTM networks may be integrated to better capture fatigue patterns over time. Additionally, deployment on embedded platforms and integration with Advanced Driver Assistance Systems (ADAS) can enable real-world in-vehicle implementation. Future extensions may also include multi-modal fatigue detection by combining facial cues with head pose or vehicle behavior, making the system more scalable and effective in real-world environments.

VII. References

[1] V. Vijaypriya and M. Uma, "Facial Feature-Based Drowsiness Detection with Multi-Scale Convolutional Neural Network," IEEE Access, vol. 11, pp. 64286-64297, 2023, doi: 10.1109/ACCESS.2023.3288008.
[2] W. Li, "Driver Fatigue Detection Based on Deep Learning Facial Features," International Journal of Transportation Science and Technology, vol. 13, no. 2, pp. 104–112, 2024.
[3] A. Majeed, S. Javaid, M. Ilyas, and M. Abid, "Real-Time Drowsiness Detection Using Deep Learning and Transfer Learning Techniques," IEEE Access, vol. 12, pp. 13245–13256, 2024.
[4] E. Çivik and H. Yüzgeç, "Real-Time Drowsiness Detection Based on CNN Using Nvidia Jetson Nano," 2023 International Conference on Artificial Intelligence and Data Science (AiDAS), Antalya, Turkey, 2023, pp. 57–62.
[5] L. Qu, T. Han, and Y. Zhang, "Driver Fatigue Detection Based on Multi-Attention Fusion Network," Applied Sciences, vol. 13, no. 6, pp. 3332, Mar. 2023.
[6] M. Xiao, F. Zhang, and X. Zhou, "FFF-CNN: A Facial Feature Fusion CNN for Driver Drowsiness Detection," Sensors, vol. 23, no. 5, pp. 2451, 2023.
[7] A. Kielty, R. Alam, and M. Aslam, "Yawn Detection Using Neuromorphic Vision and Deep Learning," Pattern Recognition Letters, vol. 164, pp. 50–58, 2023.
[8] S. S. Kang and V. Karuppusamy, "A Comprehensive Review on Fatigue Detection Using Facial and Sensor-Based Cues," Sensors, vol. 25, no. 1, pp. 50–74, Jan. 2025.
[9] S. Park and D. Kim, "Deep Learning-Based Driver Drowsiness Detection Using Full Facial Analysis," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 11, pp. 9876–9888, Nov. 2022.
[10] H. Alameen and G. Alhothali, "A Temporal CNN-LSTM Approach for Real-Time Driver Drowsiness Detection," 2022 11th International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, Malaysia, pp. 102–107.