International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Intelligent Vision and Driver Monitoring System

Jayashri Waman[1], Shubham Rane[2], Anuj Bhalerao[3], Raj Sonavane[4], Sandesh Darade[5]

[1]Professor,Department of Computer Engineering Ajeenkya D. Y. Patil School of Engineering
Pune, Maharashtra, India

[2,3,4,5] Department of Computer Engineering Ajeenkya D. Y. Patil School of Engineering
Pune, Maharashtra, India

ABSTRACT- ***Road accidents caused by driver fatigue, drowsiness, distraction, and inattentive behavior continue to be a major concern in modern transportation systems. This paper presents an Intelligent Vision and Driver Monitoring System that utilizes Computer Vision and Artificial Intelligence techniques to continuously monitor driver activities in real time. The proposed framework integrates MediaPipe Face Mesh, Eye Aspect Ratio (EAR) analysis, YOLOv8-based mobile phone detection, facial authentication, GPS tracking, emergency alert generation, trip logging, and driver analytics within a unified platform. The system analyzes facial landmarks to detect drowsiness, yawning, and abnormal head movements, while YOLOv8 is employed to identify mobile phone usage and distracted driving behavior. Whenever unsafe conditions are detected, voice-based warnings and emergency alerts are generated to regain driver attention. In critical situations, the system retrieves the driver's GPS location and supports emergency communication for rapid assistance. Experimental evaluation demonstrated reliable real-time performance with an overall system accuracy of approximately 96.0% under different operating conditions. The proposed solution operates using a standard webcam without requiring expensive hardware sensors, making it cost-effective, scalable, and suitable for deployment in personal vehicles, commercial transportation, and intelligent mobility applications.***

Index Terms— Artificial Intelligence (AI), Computer Vision, Convolutional Neural Network (CNN), Driver Monitoring, YOLO Object Dectection, MediaPipe Face Mesh Detection ,Whatsapp Emergency Alert , Eye Aspect Ratio, GPS Tracking

I. INTRODUCTION

Driver inattention — whether caused by fatigue, drowsiness, yawning, or mobile phone usage — is among the leading preventable causes of road fatalities worldwide. Unlike mechanical failures that occur abruptly, driver fatigue develops progressively, making early detection critical. Studies consistently show that reaction time deteriorates significantly after prolonged driving, yet most vehicle systems provide no feedback to the driver until a collision is imminent. Addressing this gap through intelligent real-time monitoring is the central motivation for this work.

Early approaches to driver monitoring relied heavily on physiological instrumentation — electroencephalography (EEG), electrocardiography (ECG), and respiratory sensors — which provided high-fidelity signals of driver alertness but imposed practical constraints. Wearable electrodes are uncomfortable for long trips, require careful calibration per driver, and add significant hardware cost that limits deployment to research or fleet settings. These constraints motivated a shift toward non-contact, camera-based methods that can be deployed on any vehicle equipped with a standard webcam.

Vision-based monitoring systems analyse facial features such as eye blink rate, mouth aperture, and head orientation to infer driver state. The Eye Aspect Ratio (EAR) — a geometric measure of eye openness derived from facial landmarks — has proven particularly effective as a drowsiness indicator. When EAR falls below a threshold and remains low for a sustained duration, the probability of a microsleep event is high. Simultaneously, the rise of mobile phone use at the wheel has introduced a distinct distraction pattern that EAR alone cannot capture; object detection is required to identify handheld devices in the driver's visual field.

Existing systems typically address one of these behaviours in isolation. A system that detects drowsiness may ignore phone usage; one that detects phones may have no fatigue model. Furthermore, few systems close the safety loop by generating emergency notifications with geolocation data when the driver is critically impaired. The work presented here integrates drowsiness detection, yawning analysis, head pose monitoring, mobile phone detection, facial authentication, GPS-enabled emergency alerting, and a trip analytics dashboard into a unified, low-cost platform. All processing runs on a standard webcam feed without additional sensors.

The principal contributions of this paper are as follows: (1) a real-time drowsiness detection pipeline using MediaPipe Face Mesh and EAR analysis with dual severity thresholds; (2) YOLOv8-based mobile phone detection with multi-frame verification to suppress false positives; (3) facial recognition for driver identity authentication before trip commencement; (4) GPS-integrated WhatsApp emergency alerts dispatched automatically when critical fatigue is sustained; and (5) a driver analytics dashboard that logs trip events for longitudinal safety monitoring. Experimental evaluation across 50 real-time sessions yielded an overall system accuracy of 96.0%, demonstrating practical viability under varied lighting and driving conditions.

The major contributions of this work are:

• Real-time driver drowsiness detection using MediaPipe

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Face Mesh and EAR.
• Mobile phone usage detection using YOLOv8 object detection.
• Driver authentication using facial recognition.
• GPS-enabled emergency alert generation.
• Trip logging and analytics dashboard for driver behavior monitoring.

II .LITERATURE REVIEW

Driver monitoring research has evolved along two broad trajectories: physiological signal analysis and vision-based behavioural analysis. This section traces that evolution, identifies the capabilities and limitations of representative works, and situates the proposed system relative to the state of the art.

A. Physiological and Sensor-Based Approaches
Guede-Fernández et al. [1] demonstrated that respiratory signal patterns reliably reflect fatigue-induced changes in driver alertness, achieving accurate detection using a chest-belt sensor. While their results validated the physiological basis for drowsiness monitoring, the requirement for body-worn sensors remains a meaningful barrier to mass deployment. The system could not easily generalise to ride-hailing drivers, delivery personnel, or passenger vehicle owners without additional hardware provisioning.

B. Vision-Based Facial Analysis
You et al. [2] addressed inter-driver variability by developing a real-time drowsiness algorithm that adapts to individual blinking patterns, improving robustness across a heterogeneous driver population. Their personalised model reduced false alarms in drivers who naturally blink slowly, though the adaptive calibration phase added complexity at system initialisation.
Nguyen et al. [3] applied deep learning to extract facial landmarks and compute EAR from standard webcam footage, confirming that camera-only pipelines can match the detection reliability of sensor-dependent systems. Their study established EAR as a practical, low-overhead fatigue metric and validated its use in real driving environments.
Dasgupta and Singh [4] combined Convolutional Neural Networks with OpenCV-based preprocessing to classify driver images as alert or drowsy. Their CNN architecture delivered improved accuracy over earlier threshold-based methods but concentrated exclusively on eye-state classification, without incorporating distraction detection or any form of emergency response.

C. Illumination-Robust and Multi-Parameter Systems
Patel et al. [5] specifically targeted performance degradation under low-light conditions — a known weakness in camera-based systems — and demonstrated that their deep learning model maintained stable detection accuracy during nighttime driving scenarios. Patel and Sharma [6] extended real-time monitoring to include alert generation, closing part of the action loop but stopping short of geolocation-enabled emergency communication.
Gupta et al. [7] improved detection reliability by combining EAR with head pose estimation, recognising that a fatigued driver often exhibits both eye closure and abnormal head tilt simultaneously. Their multi-parameter approach reduced misclassification in cases where eye closure alone was ambiguous — for example, drivers wearing glasses or operating in partial shadow.

D. Scalable and Multimodal Deep Learning Methods
Mehta et al. [8] demonstrated that a purely software-based CNN pipeline — without any vehicle-embedded hardware — could detect fatigue at accuracy levels competitive with sensor-augmented systems. Their work reinforced the feasibility of smartphone or webcam-only deployment, a position the proposed system adopts and extends with additional safety modules.
Singh and Kumar [9] explored multimodal fusion by combining visual face analysis with eye-movement tracking data, achieving higher drowsiness detection accuracy than single-modality approaches. The trade-off was greater computational demand, which constrained deployment on resource-limited edge hardware.

E. Research Gap and Positioning
Table I summarises the feature coverage of reviewed systems. A clear pattern emerges: existing works optimise for one or two safety dimensions — typically drowsiness detection and alert generation — while omitting features such as mobile phone detection, driver authentication, GPS-linked emergency communication, and persistent trip analytics. None of the reviewed systems provides all these capabilities in a single, sensor-free platform.
The proposed system addresses this gap by integrating MediaPipe Face Mesh [10] for high-speed landmark extraction and YOLOv8 [11] for real-time object detection within a unified monitoring framework. By combining these components with GPS tracking, WhatsApp-based emergency communication, and a driver analytics dashboard, the system moves beyond detection alone to provide a complete driver safety ecosystem operable on commodity hardware.

III.METHODOLOGY

3.1 .Eye Aspect Ratio (EAR) Computation
The proposed system uses the Eye Aspect Ratio (EAR) technique to identify driver drowsiness by monitoring eye-opening patterns in real time. EAR is obtained by measuring the relationship between the vertical and horizontal distances of eye landmark points extracted from the driver's face. The calculated EAR value decreases when the eyes begin to close. The system continuously evaluates this value frame by frame to determine the alertness level of the driver.

The drowsiness classification logic is defined as follows:
Eye closure maintained for approximately 2 to 5 seconds is treated as a normal drowsiness condition.
Eye closure extending beyond 5 seconds is identified as critical fatigue or dangerous driver inactivity.
Once the threshold condition is reached, the monitoring framework immediately activates warning mechanisms to restore driver attention and reduce accident risk.Face Mesh Detection

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

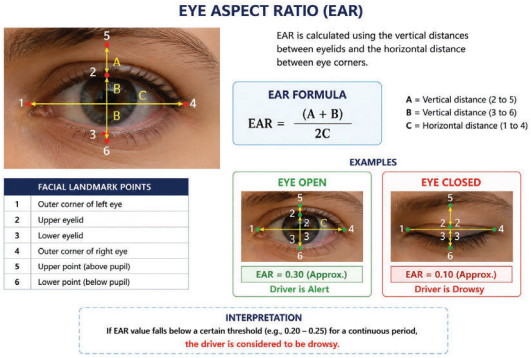

Fig. 1.Eye Aspect Ratio (EAR) Based Drowsiness Detection

The system employs the MediaPipe Face Mesh framework for high-speed facial landmark extraction. This module detects 468 facial landmark points from the driver's face in real time using a standard webcam. The extracted landmarks are utilized for multiple monitoring operations, including:

- Eye movement analysis
- Mouth opening and yawning detection
- Head orientation and movement tracking
- Driver face availability verification

The facial landmark model provides accurate feature localization even under moderate lighting variations and small head movements, enabling reliable real-time driver behavior analysis.

3.2. Mobile Phone Detection

To identify distracted driving behavior, the proposed framework integrates a YOLO-based object detection model for mobile phone recognition. Each video frame captured from the webcam is processed through the YOLO network to detect handheld devices near the driver.

The detection mechanism includes:

- Real-time object classification
- Bounding box generation
- Confidence score evaluation
- Continuous multi-frame verification

The system only confirms mobile phone usage when the object is consistently detected across multiple consecutive frames. This reduces false detection and improves monitoring stability during real-world driving conditions.

3.3. Alert and Emergency Response Mechanism

The proposed monitoring framework includes a real-time alert generation module to warn drivers whenever unsafe behavior is identified.

When abnormal conditions such as drowsiness, fatigue, distraction, or mobile phone usage are detected, the system performs the following actions:
Generates voice-based warning messages
Activates emergency siren alarms
Displays safety notifications on the dashboard
Initiates emergency communication procedures

For critical situations, the system automatically retrieves the current GPS location of the vehicle and attaches it to the emergency alert message. The location-enabled notification canthen be transmitted to emergency contacts through WhatsApp-based communication services, enabling rapid assistance and improved emergency response.

3.4. System Architecture

The proposed Intelligent Vision and Driver Monitoring System is designed as a multi-module framework that continuously observes driver activities and responds to unsafe driving conditions in real time. The architecture combines computer vision, facial landmark analysis, object detection, location tracking, and emergency communication within a single platform.
The operation of the system begins with a webcam that captures live video frames of the driver. These frames are processed using OpenCV for image acquisition and preprocessing. The processed frames are then forwarded to the MediaPipe Face Mesh module, which extracts facial landmarks required for behavioral analysis.
Using the extracted facial landmarks, the Eye Aspect Ratio (EAR) is calculated continuously to monitor eye-opening patterns. A significant reduction in EAR over a predefined period indicates driver drowsiness. The same facial landmarks are also utilized for yawning detection and head movement analysis to identify fatigue-related activities and inattentive behavior.
To detect driver distraction, the system employs a pre-trained YOLO object detection model. The model analyzes video frames and identifies the presence of mobile phones near the driver. Multi-frame verification is performed to improve detection reliability and minimize false alerts.
The outputs generated from drowsiness detection, yawning analysis, head pose estimation, and mobile phone detection are processed by the Driver Status Analyzer module. This module evaluates the driver's condition and determines whether the driver is in a normal, drowsy, distracted, or critical state.
Whenever unsafe conditions are detected, the Alert Management Module immediately activates voice-based warnings and emergency siren notifications. If critical fatigue or prolonged inattentive behavior is observed, the GPS Tracking Module retrieves the current vehicle location and prepares emergency notifications

The Emergency Communication Module transmits location-enabled alerts to predefined contacts through WhatsApp messaging services. Simultaneously, all monitoring events, alert records, and trip information are stored in the Analytics Dashboard for future analysis and reporting.

The integrated architecture enables continuous monitoring, rapid alert generation, emergency response, and driver behavior analysis without requiring expensive hardware sensors. This makes the proposed framework suitable for deployment in personal vehicles, commercial

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

transportation systems, and intelligent mobility applications.

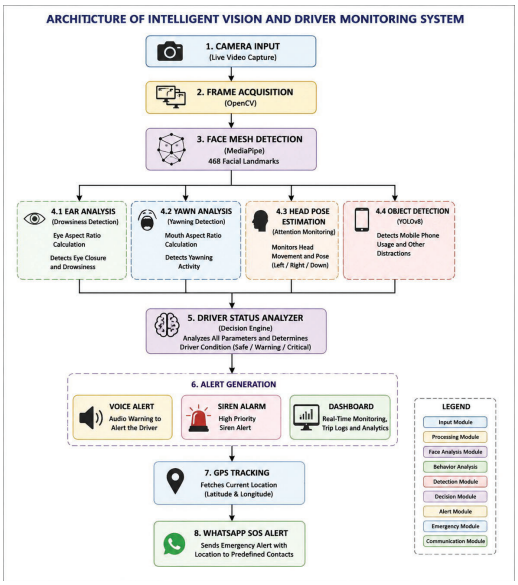

Fig. 2. Proposed System Architecture

3.5 CNN-Based Fatigue Classification

During the initial phase of development, a Convolutional Neural Network (CNN) model was implemented to classify driver facial images into alert and drowsy categories. The CNN architecture consisted of convolutional layers, max-pooling layers, and fully connected layers for automatic feature extraction and classification.

Facial images were preprocessed using resizing, normalization, and grayscale conversion before training. The objective was to automatically learn fatigue-related facial features such as eye closure patterns and facial expression changes.

Experimental evaluation showed that although the CNN model was capable of detecting fatigue conditions, its computational requirements were higher and real-time performance was less stable compared to MediaPipe Face Mesh and Eye Aspect Ratio (EAR) analysis.

Therefore, the final deployed system utilizes MediaPipe Face Mesh and EAR-based monitoring for real-time operation, while the CNN model was retained as a comparative approach for performance evaluation.

IV.RESULT AND DISCUSSION

Experimental Evaluation

The proposed Intelligent Vision and Driver Monitoring System was evaluated through real-time testing using a standard webcam under different environmental conditions. A total of 50 test sessions were conducted, including normal driving behavior, drowsiness events, yawning events, head movement scenarios, and mobile phone usage cases.

The evaluation focused on detection accuracy, alert generation speed, and system reliability. Performance was measured by comparing the detected driver state with the actual observed behavior during testing. The system demonstrated stable real-time operation and successfully generated alerts whenever unsafe driving conditions were identified.

The proposed Intelligent Vision and Driver Monitoring System was successfully implemented and tested using real-time webcam input under different operating conditions. The system continuously monitored driver activities and analyzed facial features to identify unsafe driving behavior.

The MediaPipe Face Mesh framework accurately extracted facial landmarks, enabling reliable monitoring of eye movements and facial expressions. The Eye Aspect Ratio (EAR) based approach effectively identified prolonged eye closure conditions associated with driver drowsiness. The system was able to generate warning alerts whenever the eye closure duration exceeded the predefined threshold.

Yawning detection was performed using facial landmark analysis of mouth movements. The module successfully recognized prolonged mouth opening patterns that may indicate fatigue or reduced alertness. Head movement monitoring also helped identify inattentive behavior and abnormal driver posture.

For distraction monitoring, the pre-trained YOLO object detection model successfully detected mobile phone usage near the driver. The use of continuous frame verification reduced false detections and improved overall monitoring reliability during real-time operation.

The alert generation module responded immediately whenever unsafe conditions were identified. Voice-based warnings and siren alerts were activated to regain driver attention. In critical situations, the GPS module retrieved the current location and enabled the transmission of emergency notifications through the WhatsApp alert mechanism.

The dashboard module continuously recorded monitoring events, alert history, and trip information. This functionality provided additional support for long-term driver behavior analysis and safety monitoring.

Experimental observations indicated that the proposed framework operated efficiently using a standard webcam and software-based processing without requiring specialized sensors. The integration of drowsiness detection, distraction monitoring, emergency communication, and trip logging within a single platform enhanced the overall effectiveness of the system for real-time driver safety applications.

The proposed framework successfully integrated multiple intelligent safety features such as drowsiness detection, critical fatigue monitoring, yawning analysis, head pose estimation, YOLO-based mobile phone detection, facial authentication, GPS tracking, WhatsApp emergency alerts, trip logging, and an AI-powered analytics dashboard.

By combining these modules into a single platform, the

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

system provided comprehensive driver safety monitoring without relying on expensive sensors or specialized vehicle hardware. Experimental implementation confirmed that the system can reliably detect unsafe driving behavior and generate immediate alerts under different operating conditions. The use of MediaPipe Face Mesh and OpenCV enabled accurate facial landmark extraction and behavioral analysis, while YOLO-based object detection improved distracted driving identification. Real-time warning mechanisms and emergency communication features further enhanced the effectiveness of the proposed solution.

The developed framework offers several advantages, including low implementation cost, real-time performance, scalability, and practical deployment capability for personal vehicles, commercial transportation, and intelligent mobility systems. The integration of driver analytics and trip management features also supports long-term safety analysis and transportation monitoring.Future improvements may include cloud-based data synchronization, night vision enhancement, advanced emotion recognition, voice-command integration, and IoT-enabled vehicle communication for smart transportation environments.

The system was evaluated using 50 real-time test sessions conducted under different driving conditions including normal driving, drowsiness, yawning, head movement, and mobile phone usage scenarios.

The reported accuracy values were obtained by comparing
the predicted driver state with manually observed ground truth labels during testing.

The proposed system achieved an overall accuracy of 96% during real-time testing under different driving conditions

Table.1.Model Accuracy

Parameter	Accuracy (%)
Drowsiness Detection	95.8
Yawning Detection	94.1
Head Pose Detection	95.3
Mobile Detection (YOLOv8)	96.7
Face Authentication	98.1
Overall System Accuracy	96.0

The reported accuracy values were obtained from 50 real-time test sessions conducted under different driver behavior scenarios.

Table.2.Comparing CNN And Mediapipe

Model	Accuracy
CNN	92.4%
MediaPipe + EAR	96.3%

The experimental results indicate that the MediaPipe Face Mesh and EAR-based approach achieved higher accuracy and better real-time performance than the CNN-based model. Therefore, MediaPipe + EAR was selected for deployment in the final system.

V.CONCLUSION

This work presented an Intelligent Vision and Driver Monitoring System designed to improve road safety through continuous real-time observation of driver behavior. The proposed framework successfully combined computer vision techniques, facial landmark analysis, object detection, GPS tracking, and emergency communication features into a unified monitoring platform.
The system utilized MediaPipe Face Mesh and Eye Aspect Ratio (EAR) analysis to identify driver drowsiness and fatigue-related activities. In addition, a YOLO-based object detection model enabled the recognition of mobile phone usage and distracted driving behavior. The integrated alert mechanism generated immediate warnings whenever unsafe conditions were detected, helping to improve driver awareness and reduce potential accident risks.

The inclusion of GPS tracking, WhatsApp emergency alerts, dashboard analytics, and trip logging further enhanced the practical usefulness of the proposed solution. Unlike conventional monitoring systems that depend on costly hardware sensors, the developed framework operates using a standard camera and software-based processing, making it affordable and suitable for real-world deployment.

Overall, the proposed system demonstrates the potential of intelligent vision-based monitoring solutions for enhancing transportation safety. Future improvements may include cloud-based data synchronization, advanced driver emotion recognition, night vision support, voice-assisted interaction, and integration with smart vehicle communication systems

The current implementation may experience reduced performance under extremely low-light conditions and severe face occlusions.

VI .FUTURE SCOPE

Future enhancements may include cloud-based data synchronization, advanced emotion recognition, night vision support for low-light environments, voice-command interaction, and Internet of Things (IoT) integration for smart vehicle communication. These improvements can further increase the effectiveness and scalability of intelligent driver monitoring systems.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

REFERENCES

[1] F. Guede-Fernández, M. Fernández-Chimeno ,and M. A. García-González , "Driver Drowsiness Detection Based on Respiratory Signal Analysis," IEEE Transactions on Intelligent Transportation Systems,vol. 23, no. 7,pp. 9125–9135, 2022.

[2] F. You, X. Li, Y. Gong, H. Wang, and H. Li,"A Real-Time Driving Drowsiness Detection Algorithm With Individual Differences Consideration,"IEEE Transactions on Cognitive and Development Systems,vol. 15, no. 2, pp. 284–294, 2023.

[3] Nguyen, J. Lee, and D. Kim, "Deep Learning -Based Driver Drowsiness Detection Using Facial Landmarks and Eye Aspect Ratio," Sensors, vol. 22, no. 4, p. 1645, 2022.

[4] R. Dasgupta and S. Singh, "Driver Fatigue Detection Using Convolutional Neural Networks and OpenCV," International Journal of Engineering Research & Technology (IJERT), vol. 11, no. 8, pp. 45–49, 2022.

[5] M. Patel, P. Raval, and V. Desai, "A Camera- Based Driver Drowsiness Detection System Using Deep Learning," International Conference on Smart Computing and Communications (ICSCC), IEEE, 2023.

[6] S. Patel and N. Sharma, "Real-Time Monitoring of Driver's Drowsiness Using Machine Learning and Computer Vision," IEEE International Conference on Advanced Computing (IACC),pp. 589–594, 2023.

[7] P. Gupta, A. Singh, and K. Sharma, "Driver Alertness Monitoring Using Eye Aspect Ratio and Head Pose Estimation, " Journal of Intelligent & Fuzzy Systems, vol. 45, no. 3, pp. 3771–. 3783,2023

[8] A. Mehta, S. Kulkarni, and R. Patil, "Software-Based Driver Fatigue Detection System Using Convolutional Neural Networks, " IEEE International Symposium on Artificial Intelligence for Human Safety (AIHS), pp. 12–18, 2024.

[9] M. Singh and D. Kumar, "Drowsiness Detection Using Multi-Modal Deep Learning: Fusion of Vision and Eye Movement ,Data " IEEE Access, vol. 12, pp. 45312–45325, 2024.

[10] C. Lugaresi et al.,"MediaPipe: A Framework for Building Perception Pipelines," arXiv:1906.08172, 2019

[11] G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics," GitHub Repository, 2024

(This work is licensed under a Creative Commons Attribution 4.0 International License.)