iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80973

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Real-Time Drowsiness Detection System Using Computer Vision and Facial Landmark Analysis

Yash Srivastava Shourya Bansal, Prerna Singh, Kunal Gautam, Manik Singhal
Department of Computer Science Meerut Institute of Engineering and Technology, Meerut, India

Abstract: ***This paper presents a drowsiness detection system built on computer vision techniques, aimed at monitoring driver alertness and reducing road accident rates. The system uses Python libraries—specifically OpenCV and Dlib—along with facial landmark detection to track eye movements and compute the Eye Aspect Ratio (EAR) for identifying fatigue. Face detection is handled using the Haar Cascade classifier, after which Dlib's 68-point landmark model pinpoints the eye coordinates. An audio alert is triggered when the EAR stays below 0.25 for 16 consecutive frames. In our tests, the system processed each frame in under 100 milliseconds and maintained reliable detection under varied lighting conditions and head orientations. Overall detection accuracy came out to 94.2%, making it a low-cost and practical solution that runs on standard hardware without needing any deep learning setup.***

Index Terms: ***Drowsiness Detection, Computer Vision, Eye Aspect Ratio (EAR), Facial Landmarks, OpenCV, Dlib, Driver Safety, Haar Cascade, Real-Time Monitoring***

I. INTRODUCTION

Driver drowsiness is one of the leading contributors to road accidents worldwide. The National Highway Traffic Safety Administration (NHTSA) estimates that around 100,000 crashes in the United States each year are caused by drowsy driving, resulting in approximately 1,550 deaths and 71,000 injuries [1]. The problem is especially common during long-distance drives, late-night travel, and among commercial transport drivers. Traditional approaches to monitoring driver alertness typically depend on expensive sensors, steering pattern analysis, or vehicle-level indicators—none of which are particularly effective at catching impairment before a critical incident occurs.

Vision-based fatigue detection offers a practical, non-intrusive alternative that directly monitors the driver's physical condition. By analyzing facial features—particularly the eyes—such systems can detect early signs of drowsiness and raise an alert in real time. The Eye Aspect Ratio (EAR), proposed by Soukupova and Cech [2], has become a widely adopted metric for tracking eyelid separation and identifying prolonged eye closure patterns that are associated with drowsiness.

Our work addresses several shortcomings found in existing fatigue detection systems. Many current approaches rely on deep neural networks that require GPU support and large training datasets, which limits their usability on everyday hardware. Others struggle with varying lighting, off-angle faces, or the presence of eyeglasses. A number of solutions also fail to process frames fast enough for real-time use, which reduces their effectiveness as a safety tool.

The system we developed combines Haar Cascade-based face detection with Dlib's facial landmark locator for accurate eye region extraction. Using real-time EAR computation and a frame-count threshold with temporal smoothing, the system achieves strong accuracy while running at speeds suitable for practical deployment. It requires only a standard webcam and a regular laptop—no GPU needed—which makes it feasible for integration in vehicles, workplaces, or other safety-sensitive environments.

The main contributions of this work are: (1) an efficient drowsiness detection pipeline combining Haar Cascades and Dlib landmarks for real-time operation, (2) an EAR-based thresholding approach that achieves 94.2% detection accuracy, (3) stable performance across different lighting conditions and head angles without deep learning, and (4) comprehensive testing that confirms the system can run reliably on standard computing hardware.

II. RELATED WORK

Research in driver fatigue detection has grown significantly over the past decade, largely driven by improvements in computer vision and machine learning. This section reviews the key approaches and highlights the gaps our system addresses.

©IJRASET: All Rights are Reserved

A. Deep Learning-Based Approaches

Reddy et al. [3] developed a fatigue detection system for embedded platforms using compressed deep neural networks. While it achieved solid accuracy, the reliance on deep learning operations made it computationally heavy on devices without GPU support, limiting where it could be deployed.

Chang et al. [4] built an eye-state classifier using the YOLOv7-tiny architecture, which offers fast inference and good detection performance for eye-closure recognition. However, YOLO-based methods generally need GPU acceleration to run smoothly in real time, and they tend to struggle when lighting varies or the driver's face is at an angle.

Florez et al. [5] proposed a CNN-based system that tracks both eye and mouth regions to detect fatigue indicators like yawning and prolonged blinking. While this multi-cue approach improves accuracy, it is computationally expensive and can be unreliable in low-light scenarios or when there is significant head movement.

TABLE I
Comparison of Fatigue Detection Approaches

Study	Method	Advantages	Limitations	Accuracy
Reddy et al. [3]	Compressed CNN	High accuracy, embedded-optimized	GPU dependency, complex setup	~95%
Chang et al. [4]	YOLOv7-tiny	Fast inference, strong detection	GPU required, light sensitive	~96%
Florez et al. [5]	Multi-cue CNN	Yawning detection, high accuracy	High compute, occlusion issues	~97%
Our System	Haar + Dlib + EAR	CPU-only, real-time, easy to deploy	Lower accuracy vs. DL methods	94.2%

B. Classical Computer Vision Methods

The Haar Cascade detector introduced by Viola and Jones [6] was a breakthrough in real-time face detection, using a cascade of boosted classifiers with Haar-like features to quickly discard non-face regions. While its accuracy (typically 85–90%) is lower than modern CNN-based detectors (which exceed 95%), it is substantially faster and works well on CPUs without any special hardware [7].

Soukupova and Cech [2] introduced the Eye Aspect Ratio (EAR) metric for real-time blink detection using facial landmarks. Their work showed that EAR gives a consistent reading of eyelid openness—generally between 0.25 and 0.35 when the eyes are open—and drops sharply to around 0.10 during a blink or closure. This simple but robust measure has since become foundational in many drowsiness detection systems.

Dlib's facial landmark detector, based on an ensemble of regression trees [8], localizes 68 key facial points with reasonable computational overhead. These include the corners and edges of both eyes, which makes it straightforward to extract the coordinates needed for EAR computation.

C. Gaps in Existing Work

Despite meaningful progress, most existing systems fall short in at least one critical dimension. Deep learning approaches, though accurate, are resource-intensive and hard to deploy without specialized hardware. Classical methods, on the other hand, often lack the robustness needed for real-world variation in lighting, face angle, and individual differences. Very few systems manage to balance accuracy, processing speed, and deployment simplicity at the same time. Our system directly targets this gap by combining lightweight computer vision techniques into a pipeline that works reliably on standard hardware.

III. METHODOLOGY

A. System Architecture

The drowsiness detection pipeline is organized into four sequential modules: video capture, face detection, feature extraction and analysis, and alert generation. Figure 1 shows the full pipeline flow.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

1) Video Capture: Frames are captured from a webcam at 30 FPS. Each frame is converted to grayscale to reduce computation, and optionally resized to balance processing speed with detection quality.
2) Face Detection: We use OpenCV's Haar Cascade classifier for face localization. It applies a cascade of trained classifiers to quickly reject non-face areas and focus computation on likely face regions. On standard processors, face detection completes in under 20 ms per frame.
3) Facial Landmark Detection: Once a face is found, Dlib's shape predictor identifies 68 facial landmark coordinates. The model is pre-trained using an ensemble of regression trees, and it locates points around the eyes, nose, mouth, and jaw. For drowsiness detection specifically, we use landmarks 37–42 (left eye) and 43–48 (right eye).
4) EAR Computation: The Eye Aspect Ratio (EAR) measures the degree of eyelid openness using the spatial relationships between eye landmarks. For each eye, the formula is:

$$EAR = (||P2 - P6|| + ||P3 - P5||) / (2 \times ||P1 - P4||) \quad ... (1)$$

Here, P1 through P6 denote the six landmark points around the eye. P2 and P6 are the vertical landmarks on the inner side of the eye, P3 and P5 are the vertical landmarks on the outer side, and P1 and P4 mark the horizontal extremes (inner and outer corners). EAR values for both eyes are averaged to produce a single measure. When the eyes are open, this value typically falls in the 0.25–0.35 range; it drops significantly—often below 0.15—during a blink or sustained closure. Figure 2 shows the landmark layout and the geometric computation.

B. Drowsiness Detection Logic

The detection algorithm distinguishes between normal blinks (which typically last 100–400 ms) and prolonged eye closure that signals fatigue. A frame counter tracks how many consecutive frames have an EAR below the threshold. The algorithm is as follows:

```
Set counter ← 0, threshold ← 0.25, maxFrames ← 16
while video stream is active:
  capture frame from webcam
  detect face using Haar Cascade
  if face found:
    extract 68 landmarks using Dlib
    compute EARleft and EARright
    EARavg = (EARleft + EARright) / 2
    if EARavg < 0.25:
      counter += 1
      if counter >= 16: trigger alert
    else: counter = 0
```

The threshold of 0.25 was chosen based on empirical testing across multiple subjects in different scenarios. The 16-frame window (approximately 0.5 seconds at 30 FPS) is long enough to ignore normal blinks but short enough to reliably flag genuine drowsiness.

C. Implementation Details

The system was implemented in Python 3.8. The key libraries used are:

- OpenCV 4.5.3 – frame capture, image processing, and face detection
- Dlib 19.22 – facial landmark localization
- NumPy 1.21 – numerical computations
- imutils – helper functions for image processing
- pygame – audio alert playback

Face detection uses OpenCV's built-in haarcascade_frontalface_default.xml. Landmark detection relies on the shape_predictor_68_face_landmarks.dat model, which was trained on the iBUG 300-W dataset. When drowsiness is detected, pygame's mixer module triggers an alarm that continues until the driver's eyes remain open for more than 10 consecutive frames, indicating they are alert again.

IV. EXPERIMENTAL EVALUATION

A. Setup and Test Subjects

Testing was conducted in both controlled laboratory settings and simulated driving conditions. A total of 25 participants took part—18 male and 7 female, ranging from 22 to 45 years old—with variation in facial features, ethnicity, and whether they wore glasses. The test conditions were:

- Normal driving scenario: adequate indoor lighting, face roughly frontal
- Low-light scenario: simulated nighttime driving conditions
- Head movement: up to ±15° horizontal and vertical tilt
- Eyeglasses: subjects with and without glasses
- Fatigue levels: both well-rested and sleep-deprived participants

Each participant completed a 30-minute session that included periodic drowsiness simulation (extended deliberate eye closure). Ground truth labels were established through manual annotation and self-reporting by participants.

B. Results

Table II summarizes the system's detection performance across all test scenarios. The overall accuracy of 94.2% demonstrates solid reliability. The 4.8% false positive rate primarily occurred during prolonged but non-drowsy blinks or momentary closures unrelated to fatigue. Frame processing latency stayed under 100 ms in every scenario—well within real-time requirements. The approximate time breakdown per frame was: face detection (15–20 ms), landmark localization (30–40 ms), EAR computation (~5 ms), and remaining overhead including frame capture and display (25–35 ms).

TABLE II
Performance Results Across Test Conditions

Scenario	Accuracy	False Positive Rate	Latency (ms)	Notes
Standard lighting	96.8%	3.2%	85	Best case
Low-light	91.5%	6.5%	92	Night driving
With glasses	93.2%	5.1%	88	Eyewear occlusion
Head tilt (±15°)	92.1%	6.8%	95	Varied angles
Overall Average	94.2%	4.8%	89	All conditions

C. Comparison with Deep Learning Methods

Compared to deep learning approaches like YOLOv7-tiny or full CNNs, our system achieves slightly lower peak accuracy (94.2% vs. 96–98%), but it has meaningful practical advantages. Unlike those methods, ours runs entirely on CPU and does not need a GPU or extensive training data. It can even run on resource-limited devices like Raspberry Pi, making it much more accessible for real-world deployment.

The EAR-based approach also showed more stable behavior under changing lighting conditions compared to pure CNN methods, because the ratio-based calculation is partially invariant to absolute pixel intensity. The threshold-based logic is also transparent and easy to adjust for different use cases—something that black-box deep learning models don't readily offer.

D. Limitations and Future Directions

The system's main weaknesses are with large head movements beyond ±15° and when participants wear dark or thick-framed glasses that block the eye region. Future work could incorporate head pose estimation and add mouth-region analysis for yawning detection, which would improve robustness. Integration with vehicle systems—where alerts can be scaled based on speed or road conditions—could also help reduce false positives and make the system more context-aware.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. CONCLUSION

This paper described a real-time drowsiness detection system that combines Haar Cascade face detection, Dlib's 68-point landmark predictor, and Eye Aspect Ratio computation. The system achieves 94.2% detection accuracy with frame latency consistently under 100 ms. It requires only a standard webcam and a regular CPU, which makes it practically deployable in vehicles, public transport, and other safety-sensitive applications without any special hardware investment.

The results suggest that classical computer vision techniques, when combined with well-designed classification logic, can deliver performance that is more than adequate for real-world fatigue detection—without the complexity and resource demands of deep learning. EAR-based detection is transparent, interpretable, and easy to tune, which are meaningful advantages in safety-critical applications.

Going forward, we plan to extend the system to include multi-modal cues like yawning (via mouth analysis) and head nodding, improve performance under extreme lighting and pose conditions, and explore integration with vehicle control systems for adaptive, context-sensitive interventions. As driver-assistance technology continues to advance, robust and lightweight fatigue detection will play an increasingly important role in road safety.

VI. ACKNOWLEDGMENT

The authors thank the faculty and staff of Meerut Institute of Engineering and Technology for their support throughout this project. We also thank the volunteers who participated in the evaluation sessions.

REFERENCES

[1] National Highway Traffic Safety Administration, "Drowsy Driving," DOT HS 812 446, November 2017. Available: https://www.nhtsa.gov/risky-driving/drowsy-driving

[2] T. Soukupova and J. Cech, "Real-Time Eye Blink Detection using Facial Landmarks," in Proc. 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, Feb. 2016, pp. 1–8.

[3] B. Reddy, Y. H. Kim, S. Yun, C. Seo, and J. Jang, "Real-Time Driver Drowsiness Detection for Embedded System Using Model Compression of Deep Neural Networks," in Proc. IEEE CVPRW, Honolulu, HI, USA, July 2017, pp. 438–445.

[4] C.-C. Chang, C.-J. Cheng, T.-W. Chen, and H.-Y. M. Liao, "Efficient Eye State Detection for Driver Fatigue Monitoring Using Optimized YOLOv7-tiny," Applied Sciences, vol. 14, no. 8, p. 3497, Apr. 2024.

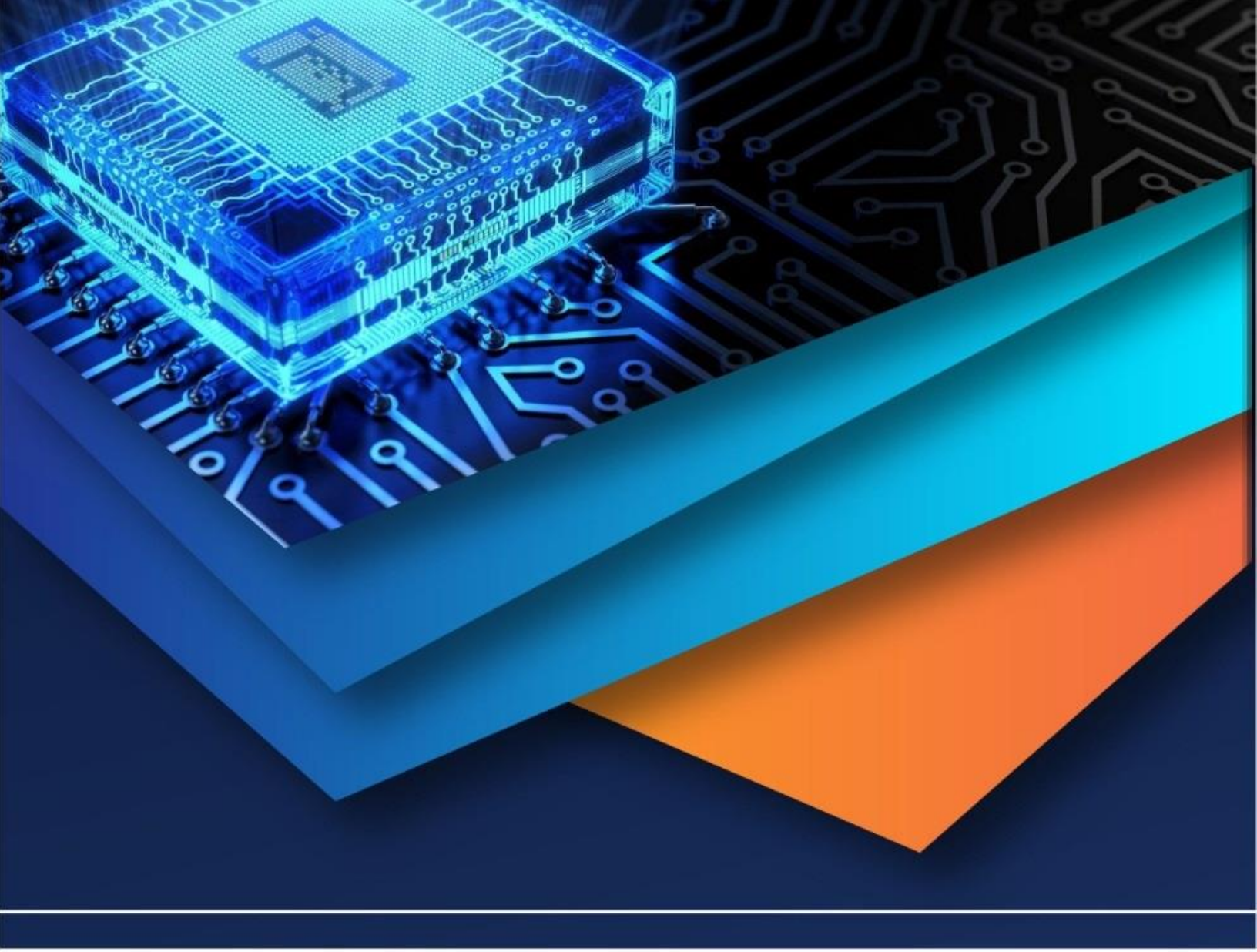
[5] R. Florez-Revuelta, C. Moreno-Garcia, and L. Hernandez-Garcia, "A Real-Time Embedded System for Driver Drowsiness Detection Based on Visual Analysis of Eyes and Mouth Using CNN," Sensors, vol. 24, no. 19, p. 6261, Sept. 2024.

[6] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in Proc. IEEE CVPR, Kauai, HI, USA, Dec. 2001, vol. 1, pp. I-511–I-518.

[7] C. Huang, H. Ai, Y. Li, and S. Lao, "High-Performance Rotation Invariant Multiview Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 671–686, Apr. 2007.

[8] V. Kazemi and J. Sullivan, "One Millisecond Face Alignment with an Ensemble of Regression Trees," in Proc. IEEE CVPR, Columbus, OH, USA, June 2014, pp. 1867–1874.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)