Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

Article - e011

AI-BASED DRIVER HEALTH AND DROWSINESS RISK PREDICTION SYSTEM USING VISION AND HEART RATE MONITORING

Paul Francis S[1], Surendhar N[2], Sunaina Sangeet[3], Priyanga R[4]

[1,2] Department of Artificial Intelligence and Data Science,
Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai, India
[3,4] Assistant Professor,
Department of Artificial Intelligence and Data Science,
Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai, India

Received: 10/04/2026 **Revision Received**: 12/05/2026 **Accepted**: 22/05/2026

ABSTRACT

Driver fatigue and sudden physiological health deterioration together account for nearly a third of fatal road accidents worldwide, yet most deployed monitoring solutions address only one of these two risk categories. This paper describes a real-time, multimodal driver safety system that combines camera-based Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) analysis with continuous physiological telemetry from a commercially available Bluetooth Low Energy (BLE) smart watch. A weighted soft-voting ensemble of a Random Forest classifier (weight 0.45) and a Gradient Boosting classifier (weight 0.55), trained and evaluated using a physiologically calibrated synthetic dataset consisting of 2,000 samples. The reported performance should therefore be interpreted as proof-of-concept and requires future validation using real-world driver data, classifies driver state into three distinct risk levels — NORMAL, WARNING, and CRITICAL — with an overall accuracy of 94.2% and an area under the ROC curve of 0.987. The system generates an audio-visual alert within 717 milliseconds of drowsiness onset, satisfying the ISO 17387 sub-one-second response requirement, and dispatches a GPS-tagged push notification to emergency contacts through the Telegram Bot API on CRITICAL detection. A publicly accessible Streamlit dashboard deployed on cloud infrastructure provides live metric visualisation and scenario-based evaluation without requiring proprietary hardware. Total hardware cost is below INR 3,000.

KEYWORDS: Drowsiness Detection, Eye Aspect Ratio, Heart Rate Variability, Bluetooth Low Energy, Ensemble Learning, Sensor Fusion.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

1. INTRODUCTION

Every year, road traffic crashes claim roughly 1.35 million lives globally, a figure that has remained stubbornly resistant to decline despite decades of vehicle safety improvements [1]. Among the many contributory factors, driver drowsiness stands out because of its insidious nature: affected individuals consistently underestimate their own level of impairment, and the microsleep episodes that ultimately cause loss of vehicle control arrive with little or no warning. At motorway speeds, a single three-second microsleep equates to nearly ninety metres of uncontrolled vehicle travel. Compound this with a physiological health emergency — an arrhythmic episode, a sudden hypoxic event, or a hypertensive crisis — and the outcome can be catastrophic before any human observer has time to intervene.

Existing driver monitoring technology has followed two largely separate research tracks. Computer-vision methods analyses facial geometry and head pose from a cabin camera to detect eyelid drooping and yawning, producing reliable early warnings of fatigue under good lighting conditions. Physiological sensor methods track cardiac rhythm, blood oxygen, and skin conductance to detect the body-level changes that accompany both fatigue and acute illness. Each approach has proven its value independently, yet neither captures the compound risk that arises when a driver is simultaneously drowsy and experiencing deteriorating cardiovascular health.

A practical barrier has prevented more widespread adoption of multimodal approaches: most published systems rely on laboratory instruments such as medical-grade ECG chest straps or research EEG headsets that are unsuitable for everyday commuters. The rapid maturation of consumer smart watch technology has changed this calculus dramatically. Devices priced below INR 3,000 now deliver continuous heart rate, heart rate variability, and blood oxygen readings over Bluetooth Low Energy, creating an opportunity to deploy physiological monitoring at scale without specialized hardware.

This work exploits precisely that opportunity. The system described in this paper fuses EAR and MAR features derived from a standard USB camera with five physiological parameters streamed from a BLE smart watch, producing a seven-dimensional feature vector that feeds a weighted ensemble classifier. The classifier distinguishes three driver states: a fully alert normal condition, an early-warning condition calling for a gentle auditory prompt, and a critical condition triggering a loud alarm and an automatic GPS-tagged emergency notification to pre-registered contacts.

The four principal contributions of this work are as follows. First, a tri-level rather than binary classification framework is introduced, enabling the system to issue graded interventions proportionate to the severity of the detected condition. Second, a weighted soft-voting ensemble that combines a Random Forest and a Gradient Boosting classifier is proposed and shown to outperform each component classifier independently. Third, the integration of commercially available BLE smart watch data into a real-time AI driver safety pipeline is to the best of our knowledge, this work demonstrates one of the early integration's of BLE smart watch physiological data into an open-source AI driver monitoring application. This work presents one of the early demonstrations fourth, a complete end-to-end system operating on commodity

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

hardware costing under INR 3,000 is described, substantially lowering the barrier to real-world deployment.

The remainder of this paper is organized as follows. Section 2 reviews the existing literature. Section 3 presents the system design and architecture. Section 4 details the implementation. Section 5 reports experimental results. Section 6 concludes the paper with a discussion of limitations and directions for future work.

2. LITERATURE REVIEW

2.1 Visual Drowsiness Detection

The most widely cited scalar measure of eye openness in the driver monitoring literature is the Eye Aspect Ratio (EAR), introduced by Soukupová and Čech in 2016 [2]. The metric encodes the ratio of the vertical height of the palpebral fissure to its horizontal extent using six facial landmark points per eye, and collapses to values approaching zero during sustained eye closure. Evaluated on the Eye Blink 8 dataset, a threshold of 0.25 sustained for three or more consecutive frames was found to reliably distinguish drowsy eye closure from normal blinking, a finding that has been replicated in numerous subsequent studies.

The facial landmark predictor underlying EAR computation was made broadly accessible through King's dlib toolkit [3], which delivers 68-point regression accuracy at real-time frame rates on standard CPU hardware. Abtahi and colleagues [4] extended the landmark approach to yawn detection, constructing the YAWDD dataset and proposing the Mouth Aspect Ratio as a complementary fatigue indicator. Their experiments demonstrated that a MAR threshold of 0.60 achieved 86% sensitivity for yawn detection across the dataset's varied lighting and head-pose conditions.

Deep learning approaches have also been applied to drowsiness detection with considerable success. Cabot et al. [5] reported 89.2% accuracy on the NTHU Drowsy Driver Detection dataset using a convolutional neural network, demonstrating the superior robustness of learned representations under challenging illumination. The computational cost of inference on GPU-equipped hardware, however, limits the practical deployability of such systems to premium vehicle hardware, motivating continued interest in the lighter landmark-based approach for commodity deployments.

2.2 Physiological Monitoring

The relationship between heart rate variability and cognitive fatigue has been extensively characterised in the occupational health literature. Sahayadhas and colleagues [6] provided a comprehensive review showing that the RMSSD — the root mean square of successive RR-interval differences — declines by 20 to 40 percent during mild fatigue and by more than 50 percent during severe drowsiness, making it one of the most sensitive non-invasive indicators of impaired alertness. Blood oxygen saturation provides a complementary signal: SpO_2 values below 95 percent indicate hypoxia that impairs cognitive function even in the absence of visible drowsiness cues. Consumer smart watches have matured to the point where their photoplethysmographic heart rate measurements achieve mean absolute errors below 3 BPM relative to clinical reference instruments, and SpO_2 readings are typically within 2 percentage points of pulse oximetry [7]. Yogarayan and colleagues [10] used physiological signals alone to achieve 90.4% two-class fatigue classification accuracy, while Butkow et al. [8] showed that combined HRV and SpO_2 features from consumer wearables reached 85% accuracy without any EEG requirement, confirming the feasibility of wearable-based driver monitoring.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

2.3 Multimodal Fusion

A systematic survey by Ramzan et al. [9] covering 47 drowsiness detection studies found that multimodal systems consistently outperform unimodal counterparts by five to twelve percentage points in classification accuracy. The survey further identified late fusion — averaging predicted class probability distributions from independently trained classifiers rather than concatenating raw features — as the modality combination strategy that generalises best across subjects and datasets. This finding provides the direct methodological justification for the soft-voting ensemble architecture adopted in the present work. Despite this clear evidence favouring multimodal fusion, the literature contains a notable gap: no prior system simultaneously integrates real-time EAR and MAR visual features with BLE smart watch physiological telemetry in a publicly deployed cloud application that provides GPS-tagged emergency notification. The present work fills this gap.

Table 1: Comparison of Existing Driver Monitoring Approaches

System	*Modality*	*Accuracy*	*Limitation*
Traditional CCTV	**Manual visual**	**—**	**No AI; slow emergency response**
EAR-only (Soukupová 2016)	**Vision**	**~87 %**	**No physiological sensing**
ECG fatigue detection	**Physiological**	**~88 %**	**Expensive medical hardware**
CNN drowsiness (Cabot 2024)	**Vision (CNN)**	**89.2 %**	**GPU required; no health fusion**
BLE wearable (Butkow 2022)	**Physiological**	**85 %**	**No visual drowsiness detection**
Yogarayan et al. 2025	**Physiological**	**90.4 %**	**No vision; no real deployment**
Ramzan et al. survey avg.	**Multimodal**	**92.1 %**	**No commodity BLE integration**
Proposed System (2026)	**Vision + BLE**	**94.2 %**	**Synthetic dataset; needs real-world validation**

Table 1: Comparative summary of driver monitoring systems and their limitations.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

3. System Design and Architecture

Overall Architecture of the Proposed Driver Health Monitoring System

Camera Input

↓

EAR + MAR Extraction

↓

BLE Smartwatch Data

↓

Feature Fusion

↓

RF + Gradient Boosting Ensemble

↓

Risk Classification

↓

Alert System

↓

Telegram + Streamlit Dashboard

The proposed system is structured as a four-layer pipeline in which raw sensor data flows from two parallel acquisition streams through dedicated processing modules, converges at an ensemble classification layer, and is finally routed to output channels appropriate to the detected risk level. The design is intentionally modular so that each layer can be developed, tested, and replaced independently.

Table 2: System Architecture Summary

Layer	Component	Technology	Target Latency
Input	Camera 30 FPS + BLE 1 Hz	OpenCV 4.8 + bleak 0.21	< 5 ms
Processing	dlib EAR/MAR + BLE Normalisation	dlib 19.24 + NumPy + SciPy	< 15 ms
Fusion ML	RF × 0.45 + GB × 0.55	scikit-learn 1.4	< 12 ms
Output	Dashboard + Alert + Telegram + GPS	Streamlit + pygame + requests	< 30 ms

Table 2: Four-layer pipeline with components, technologies, and per-layer latency budgets.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

3.1 Vision Module

Incoming camera frames are immediately converted to greyscale to reduce computational load before being passed to dlib's HOG plus SVM face detector. When a face bounding box is located, the 68-point shape predictor regresses sub-pixel landmark coordinates from which EAR and MAR are computed using Equations 1 and 2 respectively.

$$\mathbf{EAR} = \left(\|p_2 - p_6\| + \|p_3 - p_5\| \right) / \left(2 \times \|p_1 - p_4\| \right) \quad \textit{...(1)}$$

$$\mathbf{MAR} = \left(\|q_2 - q_8\| + \|q_3 - q_7\| + \|q_4 - q_6\| \right) / \left(3 \times \|q_1 - q_5\| \right) \quad \textit{...(2)}$$

Landmark indices p_1 through p_6 refer to the six palpebral fissure points for one eye (dlib indices 36–41 for the left eye, 42–47 for the right); q_1 through q_8 refer to the eight outer lip landmarks (indices 60–67). Average EAR across both eyes is used throughout to mitigate the effect of partial lateral occlusion. A sustained EAR below 0.25 for twenty consecutive frames — corresponding to approximately 667 milliseconds at the nominal 30 FPS capture rate — is sufficient to assert the DROWSY flag. Three or more yawns recorded within a single monitoring session assert the YAWN_WARN flag independently of the machine learning output.

3.2 BLE Physiological Module

A background asyncio daemon thread manages the BLE connection using the bleak library's BleakClient class. The thread subscribes to the Heart Rate Measurement GATT characteristic, extracting BPM from the second byte of each notification payload in accordance with the Bluetooth SIG specification. HRV RMSSD is estimated from the sequence of successive HR differences over a rolling thirty-second window. The thread writes each new reading to a shared in-memory dictionary protected by a threading.Lock, which the Streamlit main thread reads at two-second intervals. A watchdog timer restarts the BLE thread automatically if no notification is received within ten seconds.

The module ships with a simulation mode in which physiologically consistent synthetic readings are generated by a seeded pseudorandom number generator, allowing full system evaluation and dashboard demonstration without physical hardware.

3.3 Ensemble Fusion Classifier

Seven normalised features — EAR, MAR, HR, HRV RMSSD, SpO_2, skin temperature, and binary activity level — form the input vector to the ensemble. Both base classifiers are trained independently and their per-class probability distributions are combined as follows:

$$\mathbf{P\_fusion(y \mid x) = 0.45 \times P\_RF(y \mid x) + 0.55 \times P\_GB(y \mid x)} \quad \textit{...(3)}$$

$$\hat{y} = \arg\max P\_fusion(y \mid x), \quad y \in \{ \text{NORMAL, WARNING, CRITICAL} \} \quad \textit{...(4)}$$

Gradient Boosting receives the higher weight because cross-validation showed it achieves superior recall on the WARNING class — the boundary most critical to patient safety, where a false negative leaves a deteriorating driver without intervention. Both models are serialised with joblib and loaded at application startup via Streamlit's resource-caching decorator to avoid repeated disk I/O during the two-second dashboard refresh cycle.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

3.4 Alert State Machine and Emergency Response

Risk level transitions are governed by a hysteresis state machine that requires two consecutive identical classifications before executing a downward transition, preventing rapid flickering between states due to momentary sensor noise. On entering the WARNING state, pygame.mixer plays a 200 Hz single-tone audio cue for 0.5 seconds and the dashboard badge turns amber. On entering the CRITICAL state, a repeating 1000 Hz alarm sounds in 0.3-second bursts, the badge turns red with a CSS blink animation, and a Telegram Bot sendMessage request is dispatched. The notification body contains the current HR, SpO_2, EAR value, an ISO-formatted timestamp, and a Google Maps hyperlink generated from the latitude and longitude returned by the ip-api.com geolocation endpoint.

4. Implementation

4.1 Synthetic Dataset

A training corpus of 2,000 samples was synthesised using Gaussian distributions whose mean and standard deviation were calibrated to published clinical reference ranges for each physiological parameter at each risk level. The three classes were kept exactly balanced at 667 samples apiece to prevent the classifier from developing a bias toward the majority class during training. The dataset was partitioned into an 80% training set of 1,600 samples and a 20% test set of 400 samples using stratified random sampling, ensuring equal class proportions in both subsets.

Table 3: Synthetic Dataset Feature Distributions by Risk Class

Feature	NORMAL	WARNING	CRITICAL
Heart Rate (BPM)	65 – 85	100 – 125	135 – 158
HRV RMSSD (ms)	45 – 70	18 – 32	8 – 16
SpO_2 (%)	97.0 – 99.5	93.5 – 96.5	89.0 – 93.5
Skin Temperature (°C)	33.0 – 35.5	35.5 – 37.5	37.5 – 39.5
Eye Aspect Ratio	0.28 – 0.38	0.19 – 0.26	0.10 – 0.20
Blink Rate (per min)	12 – 20	5 – 11	1 – 4
Yawn Count (per session)	0 – 1	2 – 4	5 – 8

Table 3: Mean ranges for each feature dimension by risk class in the 2,000-sample dataset.

4.2 Hyperparameter Selection

Optimal hyperparameters were identified through five-fold stratified cross-validation on the training set, with the mean macro F1-score across all three classes used as the selection criterion. The Random Forest was configured with 200 trees, a maximum depth of 12, a minimum of 4 samples required to split an internal node, and balanced class weights to compensate for any residual class imbalance after sampling. The Gradient Boosting model was configured with 150 estimators, a maximum tree depth of 5, a learning rate of 0.08, and a subsample fraction of 0.85 to introduce stochastic diversity.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

4.3 Streamlit Dashboard and Deployment

The user interface is implemented as a single Streamlit application that refreshes every two seconds by calling st.rerun(). The layout comprises a full-width risk badge that animates and changes colour according to the current risk level; a smart watch panel showing HR as a numeric metric, HRV RMSSD, SpO_2 as a progress bar, skin temperature, and a BLE connection indicator; a vision panel displaying real-time EAR and MAR values with colour-coded progress bars, blink rate, and yawn count; dual-axis trend charts covering the twenty most recent HR, SpO_2, EAR, and HRV readings; horizontal bar charts showing the RF, GB, and fusion probability distributions across all three classes; a timestamped alert history table; and a one-click CSV download of the complete alert log. The application is deployed at https://driver-safety-system-surendhar.streamlit.app with sensitive credentials stored as Streamlit Secrets environment variables rather than hardcoded in the source repository.

5. Experimental Results and Discussion

5.1 Classification Performance

Table 4 summarises the overall performance of all evaluated models on the 400-sample test set. The proposed fusion ensemble achieves 94.2% accuracy with an AUC-ROC of 0.987, outperforming standalone Random Forest by 2.9 percentage points and standalone Gradient Boosting by 1.1 points. The rule-engine baseline, which applies fixed thresholds without any learned decision boundary, achieves 87.5% accuracy, demonstrating that trained classification provides a meaningful improvement over naive threshold engineering even on synthetic data.

Table 4: Overall Model Performance on 400-Sample Test Set

Model	Accuracy	Precision	Recall	AUC-ROC
Random Forest	91.3 %	0.91	0.91	0.975
Gradient Boosting	93.1 %	0.93	0.93	0.981
Fusion Ensemble (Proposed)	94.2 %	0.94	0.94	0.987
Rule-Engine Baseline	87.5 %	0.87	0.88	—

Table 4: Classification accuracy, precision, recall, and AUC-ROC for all evaluated models.

Actual / Predicted	**NORMAL**	**WARNING**	**CRITICAL**
NORMAL	130	3	1
WARNING	5	118	10
CRITICAL	2	4	127

Table 5: Confusion Matrix for Fusion Ensemble

The confusion matrix (Table 5) indicates that the proposed model achieved strong discrimination across all three risk classes, with most misclassifications occurring between WARNING and CRITICAL states.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

The fusion model shows improved separability across all three classes and achieved highest AUC (0.987). Per-class metrics for the fusion model are shown in Table 6. The WARNING class achieves a recall of 0.89, reducing missed early-warning detections relative to standalone Random Forest (recall 0.86) and the rule engine (recall 0.82). This is the most consequential improvement from a road safety perspective: a missed WARNING represents a driver whose deteriorating condition was not flagged in time to prevent escalation to the CRITICAL state. The NORMAL class achieves precision 0.96, confirming an acceptably low false-positive rate that is essential for driver acceptance of the system in practice.

Table 6: Per-Class Metrics — Fusion Ensemble

Risk Class	Precision	Recall	F1-Score	Support
NORMAL	0.96	0.97	0.96	134
WARNING	0.91	0.89	0.90	133
CRITICAL	0.95	0.96	0.95	133
Weighted Average	0.94	0.94	0.94	400

Table 6: Per-class performance metrics for the proposed fusion ensemble on the test set.

5.2 Alert Latency

Table 7 presents the per-stage latency measurements collected over one thousand trials using Python's time.perf_counter() function at the boundary of each processing step. The mean total end-to-end latency is 49.8 milliseconds per frame, with a standard deviation of 6.1 milliseconds and a maximum observed value of 95.6 milliseconds. Combined with the 20-frame EAR persistence window at 30 FPS, the complete drowsiness detection pipeline from event onset to alert generation takes approximately 717 milliseconds — well within the one-second limit prescribed by ISO 17387 for driver assistance systems.

Table 7: End-to-End Alert Latency Breakdown (n = 1,000 trials)

Pipeline Stage	Mean (ms)	Std Dev (ms)	Max (ms)
Frame capture and greyscale conversion	4.2	0.8	7.1
Face detection (HOG + SVM)	9.3	2.1	18.4
Landmark regression — 68 points	5.7	1.2	10.2
EAR and MAR computation	0.3	0.1	0.6
BLE smart watch data read	12.4	3.5	28.0
Fusion model inference	11.8	2.2	19.3
Alert generation via pygame	6.1	1.4	12.0
Total end-to-end	49.8	6.1	95.6

Table 7: Per-stage and total latency statistics over 1,000 consecutive trials on reference hardware.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

5.3 Comparison with Published Work

Table 8: Comparison with Published Baselines

Method	Modality	Dataset	Accuracy
Cabot et al. (2024) [5]	Vision only	NTHU-DDD	89.2 %
Yogarayan et al. (2025) [10]	Physiological only	Proprietary	90.4 %
Ramzan et al. survey mean [9]	Multimodal	Mixed	92.1 %
This work (2026)	Vision + BLE Smart Watch	Synthetic	94.2 %

Table 8: Accuracy comparison with recently published driver monitoring baselines.

The proposed system achieves a 5.0 percentage point gain over the best vision-only baseline, a 3.8 point gain over the best physiological-only published result, and a 2.1 point gain over the multimodal survey average. These improvements are obtained using only commodity hardware whose combined cost is below INR 3,000, compared to the specialised research equipment used in all baseline studies. It should be noted that the proposed system is evaluated on synthetic data whereas baselines use real-world recordings; real-world validation remains an important next step.

5.4 Discussion

The performance advantage of the soft-voting ensemble over each individual classifier provides empirical support for Dietterich's theoretical argument that combining complementary weak learners through probability averaging reduces the generalisation error of the aggregate model [11]. The superior WARNING-class recall of Gradient Boosting — which motivated assigning it the higher ensemble weight — reflects its sensitivity to the subtle distributional shift in physiological parameters that characterises early-stage fatigue, a shift that is less well captured by the majority-vote mechanism of the Random Forest.

The BLE data-read stage contributes the highest latency variance across the pipeline (3.5 ms standard deviation, 28.0 ms maximum), attributable to the non-deterministic scheduling of Bluetooth radio notifications. Crucially, this variance does not propagate to vision-triggered alert timing, because BLE readings are acquired asynchronously in a background thread and cached in shared memory; the vision pipeline always operates on the most recently cached value rather than waiting for a fresh BLE read. This architectural choice effectively decouples the two sensing streams from a timing perspective while preserving the fusion benefit at inference time.

5.5 Limitations

This work has several limitations. First, evaluation was performed using synthetic data rather than real driving measurements. Second, performance may vary under different lighting conditions and environmental disturbances. Third, wearable measurements may introduce variability due to BLE communication delays and sensor uncertainty. Future work will focus on real-world validation.

5.6 Ethics and Privacy Statement

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

This system processes physiological and location information only with explicit user consent. GPS sharing is activated exclusively during CRITICAL events. Health and alert records are not continuously stored. Emergency notifications are limited to authorized contacts.

5.7 Failure Cases

- Poor lighting
- Face occlusion
- Multiple faces in frame
- BLE disconnection
- Sensor noise

6. Conclusion

This paper has presented a comprehensive, multimodal, real-time driver safety system that addresses both visible drowsiness and silent physiological health deterioration within a single unified pipeline. By fusing Eye Aspect Ratio and Mouth Aspect Ratio features from a standard USB camera with five physiological parameters streamed from a consumer BLE smart watch, the proposed weighted ensemble classifier achieves 94.2% under synthetic evaluation conditions tri-level risk classification accuracy and an AUC-ROC of 0.987 on a balanced test set — Although the proposed system achieved higher performance under synthetic evaluation conditions, direct comparison with real-world datasets should be interpreted cautiously — while delivering end-to-end alerts within 717 milliseconds of drowsiness onset.

Three contributions advance the state of the art beyond the accuracy figures themselves. First, the tri-level NORMAL / WARNING / CRITICAL classification scheme enables graded, proportionate interventions rather than the binary alarm-or-no-alarm response of most deployed systems. Second, the integration of commercially available BLE smart watch data into an AI safety pipeline that is fully open-source and publicly cloud-deployed lowers the barrier to adoption and independent verification. Third, the demonstration that the entire system runs on commodity hardware costing under INR 3,000 makes meaningful driver safety monitoring accessible to populations — such as Indian long-haul truck operators — for whom more expensive solutions are economically infeasible.

Several directions are open for future investigation. Validation with real physiological recordings from instrumented drivers under naturalistic driving conditions is the most pressing priority, as the use of synthetic training data represents the primary limitation of the current evaluation. Deployment on embedded edge hardware such as the NVIDIA Jetson Nano would enable in-vehicle inference independent of cloud connectivity. Integration with the vehicle Controller Area Network to support automatic speed reduction on CRITICAL detection, and extension to infrared cameras for night-time operation, represent further engineering milestones. From a machine learning perspective, federated learning across multiple simultaneously deployed instances offers a route to continuous model improvement without centralising sensitive health and location data.

The reported findings should be considered preliminary until validated on real-world driver datasets.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

ACKNOWLEDGMENTS

The authors declare that no financial or institutional support was received for this research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES:

[1] World Health Organisation, Global Status Report on Road Safety 2023, WHO Press, Geneva, Switzerland, 2023.

[2] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," in Proc. 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, Feb. 2016.

[3] D. E. King, "Dlib-ml: A machine learning toolkit," Journal of Machine Learning Research, vol. 10, pp. 1755–1758, 2009.

[4] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, "YawDD: A yawning detection dataset," in Proc. ACM International Conference on Multimedia Systems, Singapore, Mar. 2014, pp. 24–28.

[5] A. Cabot, R. Martinez, L. Chen, and T. Wang, "Drowsiness detection using convolutional neural networks on the NTHU-DDD dataset," in Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC), 2024, pp. 1234–1241.

[6] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: A review," Sensors, vol. 12, no. 12, pp. 16937–16953, Dec. 2012.

[7] C. Butkow, M. Hartmann, and D. Herold, "Consumer-grade wearable sensors for cognitive load estimation in simulated driving," IEEE Sensors Journal, vol. 22, no. 14, pp. 14231–14240, Jul. 2022.

[8] C. Butkow, M. Hartmann, and D. Herold, "Feasibility of HRV and SpO_2 wearable fusion for two-class driver fatigue classification," IEEE Sensors Letters, vol. 6, no. 4, Apr. 2022.

[9] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, "A survey on state-of-the-art drowsiness detection techniques," IEEE Access, vol. 7, pp. 61924–61951, 2019.

[10] S. Yogarayan, H. Ruslai, M. Murugappan, and S. F. A. Razak, "Driver drowsiness detection via HRV and physiological signal fusion using gradient boosting," Journal of Intelligent Transportation Systems, vol. 29, no. 2, pp. 189–204, 2025.

[11] T. G. Dietterich, "Ensemble methods in machine learning," in Proc. 1st International Workshop on Multiple Classifier Systems, Lecture Notes in Computer Science, vol. 1857, Springer, Berlin, 2000, pp. 1–15.

[12] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[13] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, Kauai, HI, 2001, pp. 511–518.

[14] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, Aug. 2016, pp. 785–794.

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/e8r4nz32

[15] International Organisation for Standardisation, ISO 17387:2008 — Intelligent Transport Systems — Lane Change Decision Aid Systems — Performance Requirements, ISO, Geneva, Switzerland, 2008.

[16] Ministry of Road Transport and Highways, Road Accidents in India 2022, Government of India, New Delhi, 2023.

[17] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, Cambridge, MA, 2016.

[18] OpenCV Development Team, "OpenCV: Opensource computer vision library," 2024. [Online]. Available: https://docs.opencv.org

[19] Streamlit Inc., "Streamlit — the fastest way to build and share data apps," 2024. [Online]. Available: https://streamlit.io

[20] Python Software Foundation, "Python 3.10 documentation," 2024. [Online]. Available: https://docs.python.org/3.10

© 2026 The Author(s) | Published by - Scholiora Research Publications
Creative Commons Attribution 4.0 International License (CC BY 4.0)