Iot-Enabled Real-Time Driver Monitoring System for Drowsiness and Alcohol Detection Using Arduino and OpenCV

Usha Rani B*, Mohammad* Imran*, S.* Nishanth*, Vikas* Naik*, and Ashok* P

Dept. of CSE (IoT & CSBT), East Point College of Engineering and Technology, Bengaluru, Karnataka, India

Abstract. Driver drowsiness and alcohol impairment account for approximately 30–35% of serious road crash incidents globally, claiming over 400,000 lives annually (World Health Organization). This paper presents a real-time, non-intrusive driver impairment detection system that concurrently monitors drowsiness through computer vision and alcohol intoxication through embedded sensing. Drowsiness detection employs the Eye Aspect Ratio (EAR) derived from dlib's 68-point facial landmark regression, achieving 92.4% accuracy (F1-score = 89.7%, ROC AUC = 0.962) across 1,443 test frames. Alcohol intoxication is continuously monitored via an MQ-3 semiconductor gas sensor calibrated to the 0.05% BAC legal threshold. A three-stage hierarchical response model escalates from auditory alerting through controlled deceleration to autonomous left-lane parking with GSM emergency notification. The complete hardware subsystem is priced at USD 62–78, demonstrating practical deployability for broad vehicular safety applications.

1 Introduction

Road traffic accidents remain a leading cause of mortality and morbidity worldwide. Despite advances in vehicle safety engineering, human factors continue to dominate crash causation. Research consistently identifies driver drowsiness and alcohol intoxication as contributing factors in approximately 20–30% of fatal collisions [1]. Fatigue-induced microsleep episodes and ethanol-impaired cognition both degrade reaction time, visual processing, and hazard-anticipation ability in ways that are often unrecognised by the driver.Existing commercial solutions address these risks in isolation: lane-departure warning systems infer drowsiness from steering patterns, while ignition interlock devices require driver-initiated breath tests only at trip start. Neither approach provides concurrent, continuous, and non-intrusive monitoring of both impairment modalities. This gap motivates the present work.The proposed system bridges this gap by integrating three complementary layers: (i) a Python/OpenCV pipeline that performs real-time facial analysis at 8–15 fps on commodity hardware; (ii) an Arduino Uno (ATmega328P) that interfaces with the MQ-3 alcohol sensor, IR-based head-position sensor, and actuation subsystem; and (iii) IoT connectivity via GSM that enables emergency alert transmission without disrupting driver interaction.The primary contributions of this work are: (i) simultaneous drowsiness and alcohol detection in a single embedded platform; (ii) an autonomous hierarchical three-stage response mechanism; and (iii) hardware cost below USD 80, enabling deployment across diverse vehicle classes.

2 Literature Survey

2.1 Physiological Signal-Based Methods

Early drowsiness detection research focused on direct neurophysiological monitoring. Electroencephalography (EEG) provides high-fidelity drowsiness signatures through theta-band power increases and alpha-spindle emergence, but headset form-factor constraints preclude practical in-vehicle deployment [2]. Ogino and Mitsukura demonstrated a compact single-channel EEG system achieving 85% binary drowsiness classification; however, participant discomfort and hygiene concerns remain barriers to adoption [3]. Photoplethysmography (PPG) offers a less invasive alternative; Lee et al. integrated a wrist-worn PPG sensor with an Android platform, achieving 82% classification accuracy across five alert levels [4].

2.2 Vehicle Dynamics Monitoring

Vehicle-based approaches infer driver state from observable driving behaviour without instrumenting the driver directly. Parameters analysed include lateral lane deviation, steering wheel reversal rate, time-to-lane-crossing (TLC), and accelerator pedal dynamics. Takei and Furukawa demonstrated a correlation between steering entropy and subjectively rated fatigue levels [5]. The fundamental limitation of these methods is confounding by road geometry, surface condition, and traffic density. Commercial ADAS systems such as Mercedes-Benz ATTENTION ASSIST combine steering analysis with trip-duration heuristics, but generate excessive false-positive alerts in urban stop-and-go environments [6].

2.3 Vision-Based Facial Analysis

Non-intrusive camera-based systems have emerged as the dominant research paradigm. The foundational metric is PERCLOS (Percentage of Eyelid Closure over the Pupil over Time), originally defined by Wierwille et al. as the proportion of time eyelid closure exceeds 80% over a sliding temporal window [7]. Daza et al. demonstrated robust PERCLOS estimation from dashboard cameras under naturalistic driving conditions [8]. The Eye Aspect Ratio (EAR), a computationally lightweight geometric descriptor derived from dlib’s 68-point facial landmark regression that directly correlates with PERCLOS without requiring temporal frame-buffer integration [10]. This formulation is adopted in the present work.

Deep CNN architectures have achieved superior accuracy 97.3% classification accuracy using ResNet-18 on the NTHU-DDD dataset—but inference requirements exceeding 2 GFLOPS per frame preclude deployment on embedded systems without dedicated neural processing units [11]. Safarov et al. proposed a real-time deep learning framework combining eye-blink analysis with YOLOv8-Pose estimation, reporting 96.1% accuracy on the DROZY dataset, but requiring GPU acceleration [1]. Bekhouche et al. employed hybrid deep feature selection from video sequences, achieving strong robustness under varied lighting at the cost of elevated computational complexity [15].

2.4 Alcohol Detection Systems

Automated breath alcohol sensing in vehicles has been explored through multiple transduction mechanisms. MQ-series semiconductor gas sensors offer sub-USD 5 price points and adequate sensitivity for BAC discrimination above 0.05%; however, cross-sensitivity to methane, LPG, and ethyl acetate requires mitigation through temperature compensation and sensor baseline calibration [16-18]. Xiaoronget et al. demonstrated MQ-3 integration with a microcontroller and IoT connectivity for remote fleet management of commercial vehicles [19-21]. Ignition interlock systems require driver-initiated breath samples only at startup and cannot detect mid-journey alcohol consumption. The proposed system continuously monitors cabin ethanol vapour, addressing this critical temporal gap.

3 System Architecture

The complete system architecture comprises two coupled subsystems operating in parallel: a software-based vision processing pipeline running on a system and an Arduino Uno microcontroller platform managing physical sensors and actuators.

3.1 Hardware Platform Overview

The hardware subsystem is built around the Arduino Uno (ATmega328P, 8-bit AVR, 16 MHz crystal oscillator, 32 KB flash, 2 KB SRAM). Arduino was selected over higher-performance alternatives (Raspberry Pi, ESP32) to minimise both cost and power consumption. The total bill-of-materials cost is estimated at USD 62–78 at single-unit retail pricing. Peripheral components are organised into three functional groups: (1) Sensing—MQ-3 alcohol gas sensor (analog, pin A0), passive infrared (PIR) sensor for head position (digital, pin D4), and an ADXL345 accelerometer (I2C); (2) Communication—GSM SIM800L module for emergency SMS and GPS NEO-6M for location logging; (3) Actuation—5V active buzzer (pin D5), relay module controlling the ignition circuit (pin D6), and a 12V DC motor via L298N H-bridge (pins D9/D10) emulating throttle reduction. Full circuit schematics detailing component wiring, resistor values, and pin assignments are provided in the supplementary materials.

3.2 Vision Subsystem

Video acquisition is performed by a Logitech C920 USB camera (1080p, 30 fps) mounted on the dashboard approximately 60–80 cm from the driver’s face. OpenCV 4.5 handles frame capture and preprocessing; dlib 19.22 provides the facial landmark predictor; and PySerial enables serial communication with the Arduino over USB-CDC. The pipeline runs in Python 3.9 on a system (Intel Core i5-8250U, 8 GB RAM), processing at 8–15 fps depending on scene complexity.

3.3 System Integration and Communication Protocol

The computer vision pipeline communicates alert states to the Arduino via USB-CDC serial at 9,600 baud. A simple ASCII state-machine protocol encodes four system states: ‘A’ (Alert—normal), ‘D’ (Drowsiness detected—Stage 1), ‘S’ (Sustained drowsiness—Stage 2), and ‘E’ (Emergency—Stage 3). State transitions are validated by a checksum byte appended to each message. The Arduino main loop polls the serial buffer at 100 Hz, executing the appropriate response protocol upon valid state transitions. This decoupled architecture ensures hardware actuation continues even if the vision pipeline experiences transient processing delays.

4 Algorithmic Methodology

4.1 Preprocessing Pipeline

Each captured frame undergoes a four-stage preprocessing pipeline. (1) Grayscale conversion reduces computational load and eliminates colour-channel noise. (2) Contrast Limited Adaptive Histogram Equalisation (CLAHE, clip limit 2.0, tile grid 8×8) normalises illumination variation—critical for nighttime and backlit conditions; the clip limit of 2.0 was empirically determined to suppress noise amplification while achieving sufficient contrast enhancement. (3) Gaussian blur (5×5 kernel, $\sigma = 1.0$) suppresses high-frequency noise. (4) Frame downsampling to 640×480 pixels reduces the computational cost of face detection without loss of landmark accuracy.

4.2 Face Detection

Face detection employs the dlib frontal-face HOG+SVM detector, which outperforms the classic Viola-Jones Haar cascade under moderate illumination variation (93.8% vs. 89.2% detection rate in our test conditions) while maintaining comparable processing speed. In single-driver deployment, the largest bounding box by area is selected as the driver face region.

4.3 Facial Landmark Regression

The dlib 68-point shape predictor regresses landmark coordinates using a cascade of gradient-boosted regression trees trained on the iBUG 300-W dataset, executing in approximately 3 ms per face on a 2.4 GHz Intel Core i5 CPU. Landmarks 37–48 localise both eyes and landmarks 49–68 localise the complete mouth contour.

4.4 Eye Aspect Ratio (EAR) Computation

The Eye Aspect Ratio quantifies eyelid aperture using six landmark points per eye: $\mathrm{EAR} = (\|p_2-p_6\| + \|p_3-p_5\|) / (2\cdot \|p_1-p_4\|)$. The bilateral EAR is the arithmetic mean of left and right eye EARs. Through empirical calibration across ten participants, a threshold of $\theta = 0.25$ was established, consistent with the literature range of 0.20–0.30 [10, 14]. A drowsiness alarm is triggered when EAR remains below θ for ≥ 3.0 seconds (approximately 45 consecutive frames at 15 fps), corresponding to a microsleep episode. The 3.0-second threshold was chosen to eliminate false alarms from natural extended blinks (typically < 0.5 s) while remaining within the 4–5 second safety-critical window identified in the literature.

4.5 Mouth Aspect Ratio (MAR) for Yawn Detection

The Mouth Aspect Ratio is computed from six outer lip landmarks using an analogous formulation to EAR. A yawn is registered when MAR exceeds 0.6 for ≥ 2 consecutive seconds. The threshold of 0.6 was calibrated empirically and is consistent with prior work reporting yawn-related MAR values in the range 0.55–0.75 [14]. Yawn events are weighted at 0.4× the drowsiness score contribution of an EAR alarm, reflecting their lower per-event diagnostic specificity.

4.6 Temporal Fusion and Scoring

A rolling 60-second fatigue score accumulates weighted evidence from EAR alarms (weight 1.0), yawn events (weight 0.4), and head-nod detections from the Arduino IR sensor (weight 0.6). The score decays exponentially with a 30-second half-life to model driver recovery. Three response thresholds partition the score space: Stage 1 (score ≥ 2.0): auditory alert only; Stage 2 (score ≥ 4.0): buzzer + speed reduction; Stage 3 (score ≥ 7.0): emergency response sequence.

5 Hardware Implementation

5.1 Alcohol Detection Circuit

The MQ-3 sensor's internal resistance R_S varies inversely with ethanol concentration. The output voltage is computed through a voltage divider with load resistance R_L = 10 kΩ. Sensor warm-up requires 30–60 seconds from cold start; the Arduino firmware enforces a mandatory warm-up period before enabling alcohol detection. The detection threshold is set at ADC value 400, corresponding to approximately 0.05% BAC (polynomial fit R^2 = 0.987).

The MQ-3 semiconductor sensor exhibits non-selective response to several reducing gases including methane, LPG, and ethyl acetate. To mitigate false positives, the system implements (i) a baseline drift correction that re-samples ambient air every 10 minutes in the absence of alert triggers, and (ii) a minimum concentration-duration criterion requiring the ADC reading to exceed threshold for ≥ 2 seconds continuously before triggering an alert. Testing in a vehicle cabin environment yielded a false-positive rate of 3.1% under these mitigation measures.

5.2 Autonomous Response Actuation

Upon Stage 3 activation, the Arduino executes a three-phase parking sequence. Phase 1 activates the buzzer in a 500 ms on/off pattern for 5 seconds. Phase 2 progressively reduces the motor PWM duty cycle from 100% to 0% over 8 seconds, simulating controlled deceleration. Phase 3 reads the ultrasonic sensor (HC-SR04) to confirm a clear left-side path before activating the left indicator relay and steering left via differential motor control. The GSM module transmits an SMS containing GPS coordinates to the registered emergency contact throughout Phase 3. The autonomous parking sequence has been validated on a laboratory test bench with a scaled model; real-vehicle validation under controlled road conditions is planned as a necessary next step.

6 Experimental Results

6.1 Dataset and Testing Protocol

Testing was conducted with ten participants (7 male, 3 female; age range 22–45 years; mean age 29.6 ± 7.2 years), each providing informed written consent. Video was recorded in three conditions: (a) normal indoor laboratory lighting (500–600 lux), (b) nighttime simulation (< 50 lux with a single forward-facing LED), and (c) overhead fluorescent lighting with prescription-glasses wearers. A total of 1,443 annotated test frames were extracted from approximately 48 minutes of recorded sessions, labelled by two independent annotators with inter-rater Cohen's κ = 0.87. Ground-truth drowsiness labels were established using self-reported Karolinska Sleepiness Scale (KSS) scores ≥ 7 in conjunction with EEG alpha-power references. The ten-participant cohort constitutes a proof-of-concept validation; a larger-scale study with 50+ participants is planned as future work.

6.2 Drowsiness Detection Performance

Table 2 presents detection performance across illumination conditions. The confusion matrix for normal lighting yields: TN = 875, FP = 68, FN = 38, TP = 462, giving 92.4% overall accuracy and 92.8% specificity. The ROC AUC under normal conditions was 0.962. Performance degraded under nighttime simulation (87.5% accuracy) and when participants wore prescription glasses (89.3% accuracy). When face masks were worn, the fallback head-nod-only detection mode yielded 74.2% accuracy with elevated false-positive rate.

Table 2. Drowsiness detection performance under varying illumination conditions.

Condition	Accuracy (%)	F1-Score (%)	FPR (%)	Specificity (%)
Normal lighting	92.4	89.7	7.1	92.8
Nighttime sim.	87.5	85.2	10.3	89.7
Prescription glasses	89.3	87.1	8.6	91.4
Face mask (fallback)	74.2	71.5	15.4	84.6

* Fallback head-nod-only mode active (face occlusion by mask).

6.3 Alcohol Detection Performance

The MQ-3 alcohol sensor demonstrated 100% detection reliability at legally intoxicated levels (BAC ≥ 0.08%) and 91.3% at borderline levels (0.05%–0.08%), with a 3.1% false-positive rate from environmental contaminants under active baseline correction. System response latency averaged 3.2 ± 0.4 seconds from drowsiness onset to alert activation, primarily attributable to the intentional 3.0-second EAR confirmation window.

6.4 Comparative Benchmarking

Table 1 contextualises the proposed system against representative approaches from the literature. The proposed system accepts a 4.9% accuracy reduction relative to deep CNN baselines (97.3%, [11]) in exchange for a greater than 6× cost reduction (USD 78 vs. USD 450+) and greater than 100× computational efficiency, enabling real-time embedded deployment without dedicated neural processing hardware. No prior single-platform work in the reviewed literature simultaneously addresses both drowsiness and alcohol detection with autonomous corrective actuation.

Table 1. Comparison of driver impairment detection approaches.

Method	Intrusiv e	Alcohol Detection	Autonomous Response	Cost
EEG-based	Yes	No	No	High
Steering-based	No	No	No	Medium
PERCLOS	No	No	No	Medium
CNN-based [11]	No	No	No	High
Proposed System	No	Yes	Yes	Low (~$78)

7 Limitations

Several limitations of the current implementation warrant acknowledgment. First, face detection performance degrades below 60% under direct sunlight causing camera saturation; this is being addressed through HDR frame compositing. Second, face occlusions from surgical masks, sunglasses, and balaclava-style head coverings render landmark-based EAR computation unreliable; a fallback head-nod-only detection mode is activated in these cases with reduced sensitivity (Table 2). Third, the MQ-3 sensor requires 30–60 seconds of warm-up from cold start, creating a vulnerability window; a parallel heated catalytic sensor is under evaluation. Fourth, the current single-camera setup does not capture gaze direction, a valuable distraction-detection cue.From a deployment perspective, the USB tether between system and Arduino creates cable management challenges in production vehicles. Wireless communication via Bluetooth or Wi-Fi Direct is the preferred pathway for commercial integration. Regulatory approval pathways for autonomous vehicle intervention systems vary significantly by jurisdiction and represent a practical barrier to commercialisation independent of technical readiness. Finally, the autonomous parking sequence has only been validated on a laboratory test bench; real-vehicle validation under controlled road conditions is a necessary next step before any clinical or commercial deployment.

8 Conclusion

This paper presented a comprehensive real-time driver impairment detection system integrating non-intrusive computer vision-based drowsiness recognition with hardware-level alcohol sensing on an Arduino Uno platform. The Eye Aspect Ratio metric, computed from dlib's 68-point facial landmark regression, provided robust microsleep detection with 92.4% accuracy and 3.2-second mean response latency under varied illumination conditions. The MQ-3-based alcohol detection subsystem demonstrated reliable discrimination at legally relevant BAC thresholds with active cross-sensitivity mitigation.

A hierarchical three-stage response model escalates from auditory alerting through controlled deceleration to autonomous left-lane parking, with concurrent GSM-based emergency notification. The system's key differentiating characteristics are the simultaneous coverage of both impairment modalities, autonomous corrective actuation, and sub-USD 80 hardware cost enabling broad accessibility.

Future work will address identified limitations through: deep learning-based eye-state classification (YOLOv8-Pose or MediaPipe Face Mesh) for improved low-light robustness; V2X communication integration for coordinated traffic management; PPG biometric fusion for multi-modal fatigue confidence scoring; and a compact embedded Linux platform (Raspberry Pi 5) to eliminate system dependency and facilitate real-vehicle validation.

Acknowledgements

The authors thank all study participants for their time and cooperation, and acknowledge the open-source contributions of the OpenCV, dlib, and Arduino communities.

References

1. F. Safarov, F. Akhmedov, A. B. Abdusalomov, R. Nasimov, and Y. I. Cho, Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety. Sensors, 23(14), 6459 (2023).

2. M. I. Basheer Ahmed et al., A Real-Time Computer Vision Based Approach to Detection and Classification of Driver Drowsiness. Big Data and Cognitive Computing, 7(1), 22 (2023).
3. Y. Albadawi, A. AlRedhaei, and M. Takruri, Real-Time Machine Learning-Based Driver Drowsiness Detection Using Visual Features. Journal of Imaging, 9(5), 91 (2023).
4. F. Majeed, U. Shafique, M. Safran, S. Alfarhood, and I. Ashraf, Detection of Drowsiness among Drivers Using Novel Deep Convolutional Neural Network Model. Sensors, 23(21), 8741 (2023).
5. A. Roy and J. Bhaduri, A Deep Learning Enabled Multi-Class Plant Disease Detection Model Based on Computer Vision. AI, 2, 413–428 (2021).
6. V. Phan, T. N. Trieu, and T. C. Phan, Driver drowsiness detection and smart alerting using deep learning and IoT. Internet of Things, 22, 100705 (2023).
7. S. E. Bekhouche, Y. Ruichek, and F. Dornaika, Driver drowsiness detection in video sequences using hybrid selection of deep features. Knowledge-Based Systems, 252, 109436 (2022).
8. J. Zafar, Arduino and MQ-3 Sensor Based Alcohol Detector. MA Robotic (2024).
9. Figaro Engineering Inc., TGS 2620 Datasheet: Sensor for Solvent Vapors. Technical Datasheet, 2019. [MQ-3 cross-sensitivity: Winsen Electronics MQ-3 product sheet.]
10. W. Xiaoronget et al., IoT-based alcohol detection for commercial vehicle fleet management using MQ-3. J. Ambient Intelligence and Humanized Computing, 11, 4315–4326, 2020.
11. Y. Albadawi, M. Takruri, and M. Awad, A review of recent developments in driver drowsiness detection systems. Sensors, 22(5), 2069 (2022).
12. A. Smolinski, P. Forczmaski, and A. Nowosielski, Processing and integration of multimodal image data supporting detection of behaviors related to reduced concentration level. Electronics, 13, 2457 (2024).
13. S. Jamil, M. Jalil Piran, and O. J. Kwon, A comprehensive survey of transformers for computer vision. Drones, 7(5), 287 (2023).
14. V. Vaithianathan, P. Raja, S. S. Sinnasamy, R. Gayathri, and B. Vasudevan, "Free space optics transmission enabled by PAM-4 signals: Performance evaluation under Indian weather conditions," *Internet Technol. Lett.*, **8**, e70004 (2025).
15. B. Ebenezer Abishek, A. Raaza, S. Ramesh, S. Jerritta, and V. Rajendran, "Circularly polarized circular slit planar antenna for vehicular satellite applications," *ACES Journal*, **34**, 1340–1345 (2019).
16. K. Kayalvizhi and S. Ramesh, "Design and analysis of reactive load dipole antenna using genetic algorithm optimization," *Appl. Comput. Electromagn. Soc. J.*, **35**, 279–287 (2020).