

Multi-Modal Edge Intelligence System for In-Cabin Driver Monitoring and Preventive Interventions

Sindhu Pazhayakandathil¹, Akshay George², Sarin Surendran Nambiar³, Sanjay Madathil Kuniyil⁴,
Sanjeev Siva Chakkath Sivadasan⁵

^{1,2,3,4,5}Department of Electronics and Communication Engineering, Government Engineering College
Wayanad, APJ Abdul kalam technological University, Kerala, India

Corresponding Author Email: sindhu.p@gmail.com

Abstract— Operator impairment remains a leading cause of fatal collisions, and modern intelligent transport safety systems rely on multi-modal analysis for real-time risk mitigation. This paper presents a multi-modal Internet of Things (IoT)-embedded system designed for real-time driver state valuation and preventive intervention. By pairing Raspberry Pi 5 with computer vision arrays and isolated physiological telemetry, the proposed work mitigates risks associated with operator fatigue, cognitive distraction, and medical emergencies. Contributions of this work include the integration of localized landmark extraction to resolve facial geometry dynamics alongside an electrically isolated AD8232 electrocardiogram (ECG) pipeline that captures cardiovascular state vectors to compute heart-rate anomalies. Furthermore, the system restricts cognitive multi-tasking by using contour parsing optimized for mobile phone detection. It enforces automated interventions via isolated pulse-width modulation deceleration and cloud-synchronized emergency notifications. Empirical validation proves that sub-second latency constraints are met, confirming the architecture's operational readiness.

Key Words: Edge Computing, Multimodal Sensor Fusion, Isolated Telemetry, Computer Vision, Internet of Things, Driver Drowsiness Detection, Intelligent Transportation Systems.

I. INTRODUCTION

Road accidents remain a major global health crisis. The World Health Organization reports about 1.35 million preventable traffic deaths every year [1]. A deep look at these numbers shows that driver drowsiness is a major cause of crashes. It causes 15% to 30% of all collisions and up to 20% of fatal crashes in developed nations [1]. A human life cannot be recovered or refilled once gone. Even in non-fatal accidents, financial damage is huge. Global losses exceed \$100 billion annually. This total includes direct costs like medical bills, emergency response, and car repairs. It also covers hidden costs like lost productivity, higher insurance rates, and long-term disability care. This huge impact shows the urgent need for new technology to prevent these accidents.

Drowsiness is complex and hard to study compared to other driver impairments. It is a slow shift along the wakefulness-to-sleep continuum. As a driver gets tired, their body and mind undergo subtle changes [2]. Brainwaves shift toward slow theta waves (4–7 Hz) and lose fast alpha waves (8–12 Hz) on electroencephalogram (EEG) trackers. Eye metrics also change. Blink durations slow down from 100–150 ms to 300–500 ms, and percentage of eyelid closure scores (PERCLOS) rise. At the same time, driver reaction times slow down by 30% to 50% compared to normal alert states [3]. Unlike alcohol limits which are easy to test in blood, drowsiness builds slowly. It varies heavily between people based on body clocks, sleep debt, and genetics. Also, tired drivers suffer from poor self-awareness and rarely realize how sleepy they are. This combination makes driving while tired very dangerous. It requires smart technology to catch the early signs of fatigue [4].

Recent changes in car technology offer great opportunities for advanced drowsiness detection [5, 6]. Modern cars now feature smart systems that combine three main tracking methods. First, computer vision setups use high-resolution, infrared dashboard cameras (1080p at 60 fps) with 68-point facial landmark trackers. Second, body sensor networks use steering wheel PPG sensors to check heart rate changes and capacitive sensors to measure skin response. Third, vehicle telemetry tracks steering wheel angles, which drift from 0.5°–1.5° up to 2°–5° when a driver gets tired, alongside lane deviations [7]. These multiple data streams are processed through hybrid models. Today, these advanced systems achieve strong, reliable tracking results on real roads [8, 9].

II. LITERATURE REVIEW

There are physiological, biological, and cognitive parameters through which driver drowsiness and distraction can be detected. Historically, systems relied on isolated tracking methodologies whereas modern designs favor integrated multi-sensor fusion. Some of the major works based on these different metrics are briefed in the following sub-sections.

A Vision-Based Drowsiness Detection

The driver's eyes and face give us an immediate indication about drowsiness. It is considered the most common way to spot drowsiness. By tracking eye states, blink frequencies, and micro-expressions, systems can catch early fatigue before performance of driver drops [10]. These measurable behavioral changes are regarded as reliable physical signs [11, 12, 13, 14]. Researchers use in-cabin cameras and computer vision to measure blink rates [15], maximum eye closures [16], and PERCLOS [3]. They also calculate the eye aspect ratio (EAR) [17], track real-time eye changes [18], map pupil drift patterns [19], and trace gaze vectors [20]. More advanced methods map high-dimensional geometric configurations using local facial landmarks [21]. They count yawning frequencies [22] and evaluate multi-axis head orientations [23] for accurate tracking.

To refine real-time spatial tracking on low-power processors, local computations are carried out by edge toolsets. They deploy lightweight 3D convolutional neural networks (3D CNNs) that consider spatial and temporal motion blocks across sliding multi-frame video arrays [24] for smooth results. Landmark setups have also advanced to single-stage networks like YOLO-Face Mark [25]. This model simultaneously localizes head posture bounding boxes and structural key points under difficult night conditions without multi-step loops.

Despite high remarkable accuracies often exceeding 95%, a large gap limits real-world deployment. Accuracies heavily rely on within-dataset or subject-dependent testing. When evaluated across independent subjects or real-world scenarios, performance estimates drop by 15% to 20%, and false-alarm metrics rise due to data shifts, ambient night glare, and safety eyewear occlusions [11]. To improve performance, standalone driver monitoring systems for mobile smart terminals to track blink, yawn, and head-tilt vectors have been developed [26].

B. Biological Signals and Body Telemetry

Biological signals give objective, real-time insight into the internal physiological state of a driver. This gives them a highly reliable modality for early fatigue and drowsiness detection [5, 27]. EEG setups monitor raw scalp electrodynamics. They chart a distinct increase in theta wave activity and a decrease in alpha wave activity to track sleep transitions before behavior degrades [28]. Because dry caps with wires cause driver discomfort, modern automotive technologies leverage non-invasive wearable arrays and steering-wheel sensors.

Electrocardiography (ECG) captures cardiac electrical waveforms to process autonomic changes in heart rhythms. Photoplethysmography (PPG) tracks blood volume fluctuations through the skin via optical sensing modules embedded in the wheel trim [29]. Both modalities compute Heart Rate Variability (HRV) feature

metrics across time and frequency. They specifically track mean interval vectors (MeanNN), standard deviations (SDNN), root mean square differences (RMSSD), and low-to-high frequency power ratios (LF/HF). These metrics can be used for drowsiness measurement as they flag parasympathetic nervous system dominance as driver fatigue sets in [30].

Electrooculography (EOG) isolates corneo-retinal voltage balances to provide rapid eye movement and blink monitoring benchmarks that remain accurate during extreme head deflections [31]. Electromyography (EMG) isolates electrical activity across facial muscle groups to identify muscle tone loss preceding microsleep cycles [7]. However, a high computational burden is often associated with accurate prediction. Embedded processing cores must remove serious motion-induced noise and muscle artifacts before performing classification. To expand tracking dimensions past raw comfort limits, contemporary studies explore multi-modal physiological fusion. A dual-sensing setup coupling of EEG and EMG signals processed via Continuous Wavelet Transform (CWT) and an optimized VGG-16 backbone was introduced [32].

C. Cognitive Distraction Telemetry Frameworks

Driver multitasking via mobile phone usage introduces a highly hazardous cognitive barrier. It degrades in-lane corrective responses by up to 40% compared to an alert baseline. Early attempts to track handheld targets depended on centralized deep learning models, such as standard YOLO networks, deployed on secondary desktop hardware panels [33]. While achieving precise object bounding, these isolated frameworks incurred excessive execution overhead and bandwidth costs. This limited their effectiveness for responsive, low-latency deployment on restricted edge microcontrollers.

Recent implementations attempt to simplify search domains by tracking localized pixel changes. However, standard grayscale thresholding matrices introduce high false-positive indices under transient environmental glare or shadows. To overcome this vulnerability without exceeding edge processing limitations, modern architectures employ dual color-space matrices (HSV and YCrCb). This approach suppresses human skin tones while highlighting target phone casing geometry [26]. Cross-referencing these segmented contours against structural facial landmark anchors establishes a robust distance validation loop (d_{prox}). This allows the edge node to distinguish brief hand adjustments from true driving hazards.

D. Vehicle Dynamics and Multi-Sensor Fusion

Vehicle-based measurements leverage the vehicle's existing embedded sensor infrastructure to detect drowsiness-induced performance degradation [34]. Microcontrollers monitor steering wheel movements via internal encoders. They alert when a normal operator's

precise micro-corrections (0.5° – 1.5° paths) break down into erratic, delayed steering wheel angle corrections (reaching 2° – 5° deflective angles) alongside unstable lane deviations [7]. While vehicle telemetry is completely non-intrusive and highly cost-effective, it operates as a late-stage indicator. It normally identifies fatigue only after operational safety compromises have already taken place [26].

To bridge this gap, modern Intelligent Transportation Systems deploy heterogeneous multi-sensor fusion architectures. Integrating early-stage physiological indicators with behavioral face-mesh matrices and delayed vehicular dynamic data streams significantly enhances system resilience [35]. These setups combine raw metrics via localized attention loops or late-fusion thresholding strategies. This workflow successfully runs deep learning classification tasks on ARM-based or embedded GPU platforms within tight 2–5W edge power boundaries.

E. Key Contributions of the Paper

An end-to-end, deployable, intelligent multi-modal Internet of Things (IoT)-embedded safety system is proposed for continuous real-time driver state evaluation and emergency interventions. The primary contributions of this paper are detailed below:

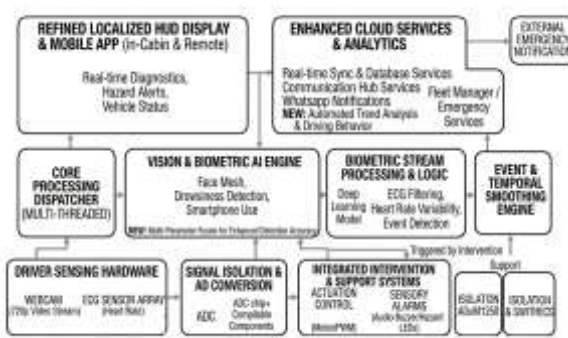


Fig - 1: Proposed Multi-Modal Edge Intelligence Framework for Driver Safety

1. Drowsiness Detection Module: Used a multi-point spatial coordinate matrix utilizing dlib's pre-trained facial landmark predictor to calculate the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) to identify eyeblinks and yawning over consecutive frames. This approach eliminates spatial tracking vulnerabilities during rapid head deflections and applies an evaluation window filter (W_E) to distinguish normal blinks from the normal microsleep events.

2. Cognitive Distraction Parsing Engine: Developed a dual-space (HSV and YCrCb) color-space transformation pipeline to isolate dark rectangular smartphone casings while suppressing dynamic driver skin tones under minimum cabin lighting. This framework uses an

oriented bounding box (OBB) constraint mechanism together with real-time Euclidean spatial proximity mapping (d_{prox}) against outer ear/jaw landmarks to confirm manual cell phone distractions within consecutive frames.

3. Isolated Cardiovascular Telemetry Core: Designed an analog processing front-end using an ADuM1250 bidirectional digital isolator to decouple the serial I2C bus. This architecture isolates the driver's skin-electrode interface from high-current actuator spikes and couples it with a digital bandpass filter (0.65 Hz to 40 Hz) and a dynamic second-derivative R-peak localization tracker to compute heart rate variability metrics which are free from motion artifacts.

4. Multi-Stage Localized Actuation and Alarm Network: Programmed a dynamic alarm controller connected via EL817 optoisolators to translate digital control commands into light pulses, protecting the processor from inductive feedback. The system manages a synchronized local warning array—combining an active 5V audio buzzer with high-luminance hazard LEDs—to apply graduated local alert patterns indicating the severity of the danger.

5. Multi-Threaded SOS Wireless Dispatch Pipeline: Implemented a communication handler that resolves network I/O bottlenecks, allowing cloud routines to run concurrently without slowing the primary vision pipelines. This engine pushes compressed telemetry matrices to Firebase Realtime Database clusters while executing automated Twilio WhatsApp SOS alerts with precise vehicle safety payloads within sub-second thresholds.

6. Graduated Vehicle Deceleration Subsystem: Designed an automated speed control layer that manages vehicle velocity by altering the pulse-width modulation (PWM) duty cycles of an L298N H-bridge driver through isolated IRFZ44N power MOSFETs switches. This setup bypasses dangerous abrupt stops by safely downscaling mechanical propulsion metrics to a controlled halt if a critical drowsiness or cardiovascular anomaly profile extends beyond the specified time limits.

III. MATHEMATICAL FORMULATIONS

The mathematical formulations for detecting the visual, cognitive, and physiological indicators are described in the following sub-sections:

A. Ocular Dynamics via Multi-Point EAR

Fatigue validation requires tracking multi-point eyelids coordinates. Let $p_i = (x_i, y_i)$ define specific 2D landmark locations on the face mesh. The system determines the Eye Aspect Ratio (EAR) by analyzing the ratio of vertical eye distances to horizontal eye width:

$$EAR = (|p_{12} - p_{16}| + |p_{13} - p_{15}|) / (2|p_{11} - p_{14}|) \quad (1)$$

where landmarks p_{11} and p_{14} mark the horizontal outer corners, while pairs (p_{12}, p_{16}) and (p_{13}, p_{15}) define the matching internal vertical coordinate pairs. The sliding evaluation metric scales over a sequential frame window W_E :

$$f(\text{Drowsy}) = 1 \text{ if } (1/W_E) * \sum I(\text{EAR}_t < \tau_E) \geq 0.85, \text{ else } 0 \quad (2)$$

Here, I is an indicator function checking if conditions are met, and τ_E represents the drowsiness threshold (empirically set to 0.2).

B. Oral Geometry Analysis (MAR)

The system separates speech-driven jaw movements from genuine fatigue yawns using the Mouth Aspect Ratio (MAR) as:

$$\text{MAR} = (|p_{22} - p_{28}| + |p_{23} - p_{27}| + |p_{24} - p_{26}|) / (2|p_{21} - p_{25}|) \quad (3)$$

where landmarks p_{21} and p_{25} define the horizontal boundaries of the lips, and the numerator terms capture vertical tracking vectors across the inner oral cavity. Yawning behavior is flagged if the ratio satisfies:

$$\int H(\text{MAR}(t) - \tau_M) dt \geq \Delta t_Y \text{ (where } \Delta t_Y = 2.3 \text{ s)} \quad (4)$$

C. Perspective Head Pose Estimation

The head tracking algorithm approximates spatial tilt by measuring the lateral displacement of the face bounding box relative to the central coordinate axis of the frame. Let W denote the absolute frame width in pixels. The central vertical axis of the sensor layout is defined as $W/2$. Given the horizontal coordinate of the face centroid x_b , the absolute pixel displacement Δx is formulated as:

$$\Delta x = |x_b - W/2| \quad (5)$$

$$\hat{x} = \Delta x / (W/2) = 2|x_b - W/2| / W \quad (6)$$

$$\theta(t) = \min(\hat{x} * 35, 45) \quad (7)$$

$$\bar{\theta}(t) = (1/M) * \sum \theta(t - i), \text{ where } M = \min(t + 1, 5) \quad (8)$$

D. Cardiovascular Waveform Processing

Digital outputs from the 16-bit ADC are converted to a time-series voltage array $V[n]$. To suppress muscular artifact noise, a digital bandpass filter extracts the cleanest signals between 0.5 Hz and 40 Hz. Real-time heart rate calculation applies a derivative thresholding window to locate consecutive R-wave peaks:

$$R_{\text{peaks}} = \{ n \mid dV[n]/dn = 0, d^2V[n]/dn^2 < 0, V[n] \geq \gamma_{\text{adapt}} \} \quad (9)$$

$$\text{BPM} = (60 * f_s) / (n_k - n_{k-1}) \quad (10)$$

E. OpenCV Smartphone Contour Detection Framework

To minimize distraction-induced collisions, a vision-based pipeline runs on the edge node to identify mobile phones within the driver's immediate peripheral zones. The algorithm relies on color segmentation, thresholding matrices, geometric constraints, and morphological filters to separate phone contours from skin tones and background noise.

IV. METHODOLOGY

The drowsiness detection system addresses the problem in a multifaceted way. The real-time drowsiness framework continuously captures in-cabin camera frames. It tracks the driver's facial landmarks to compute eye and mouth metrics. If the system detects eye closure past 0.85 seconds or continuous yawning past 2.3 seconds, a drowsiness flag is raised. If a head tilt of more than 41° is detected, another flag is set. The system parallelly separates skin colors from surrounding objects to isolate background noise, and then identifies mobile devices by validating their rectangular aspect ratios. If a mobile phone structure crosses a 150-pixel proximity threshold next to the face for over 1.3 seconds, it sets a distraction flag. In a separate pipeline, ECG is monitored through the connected electrodes, and comparing it with the driver's stored profile, any anomalies can trigger the health flag. Any flagging event will activate the hazard lights and alarms, and will cause the vehicle to decelerate immediately, making it to a complete halt. This results in an early intervention before the driver loses complete control and thus can prevent accidents. At the same time, SOS messages are sent to the stored emergency contacts, and parallelly over the cloud for the admin panel to monitor and intervene at any time. The whole system is built around a Raspberry Pi core as explained in the following section.

V. CIRCUIT TOPOLOGY AND ELECTRICAL ISOLATION

The schematic deployment implements galvanic isolation barriers and low-side switching topologies to guarantee driver safety, preserve signal integrity, and eliminate inductive feedback from high-current actuator loads.

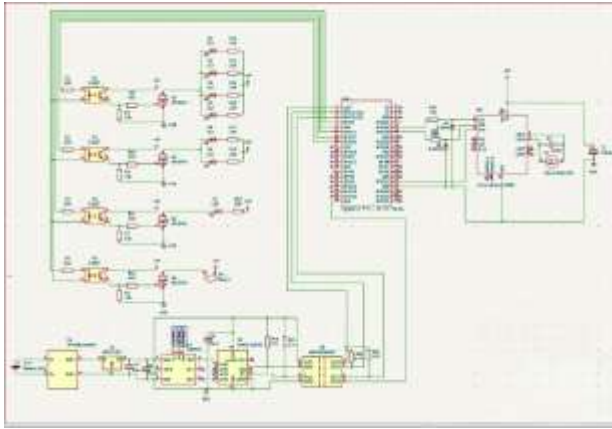


Fig - 2: Circuit Diagram

- Isolated Biometric Sensing Topology:** The physiological front-end connects skin electrodes to an AD8232 ECG core. To prevent power rail ripple from introducing noise, analog voltages pass to an ADS1115 16-bit Delta-Sigma ADC. The digital I2C bus lines (SDA/SCL) are pulled high via 4.7 k Ω resistors and decoupled through an ADuM1250 bidirectional digital isolator. The ADuM1250 utilizes micro-machined, chip-scale air-core transformers to maintain a complete galvanic boundary, preventing computing core voltage spikes from reaching the driver's skin.
- Optocoupled Low-Side Switching Array:** High-frequency digital control commands travel from Raspberry Pi GPIO pins (GPIO2, GPIO27, GPIO22, GPIO17) through 330 Ω inline current-limiting resistors to drive the infrared LEDs of EL817 optoisolators. The receiving phototransistors alter the gates of IRFZ44N N-Channel Power MOSFETs ($R_{DS(on)} \leq 17.5 \text{ m}\Omega$). Each gate includes a 430 Ω pull-up resistor paired with a 10 k Ω gate-to-source pulldown resistor to guarantee accurate switching states. This arrangement switches an active 5V audio buzzer and a high-luminance status LED cluster without common-ground interference.
- Graduated Deceleration Actuator Stage:** To perform controlled emergency braking, the system integrates an L298N Dual H-Bridge Motor Driver connected via isolated IRFZ44N power switches. Direction inputs (IN1, IN2) and the primary speed pin (EnA) connect to Raspberry Pi pins (GPIO15, GPIO18, GPIO12) via a protection network containing 330 Ω series resistors and protective Zener diodes. The driver alters the pulse-width modulation (PWM) duty cycles of an external DC propulsion motor. To buffer inductive surges and motor feedback, the output rail incorporates a 100 nF ceramic capacitor in parallel with a 1000 μF electrolytic bulk capacitor.
- Power Regulation Architecture:** The device is powered by a portable 3.7V Lithium-Ion Battery Cell regulated by an onboard TP4056 charging module. The charging stage output passes to an NCP1700T Low-Dropout (LDO) linear regulator. Paired with filtering capacitors (1 μF and 0.1 μF), it supplies a stable, noise-

free +3.3V reference voltage to the analog sensing electronics.

VII. EXPERIMENTAL RESULTS AND EVALUATION

Testing confirms that the framework successfully meets core real-time processing constraints on the Raspberry Pi 5 platform under varying environmental conditions.

A. Latency and Real-time Performance

Spatial landmark extraction loops maintain a steady processing rate between 32 and 35 frames per second (FPS) under moderate cabin illumination profiles. To validate compliance with real-time vehicular safety constraints, execution overhead profiling is conducted across discrete sub-routines:

- Face Mesh Processing:** Latency registers tightly at 27 ms.
- OpenCV Shape/Contour Parsing:** Local execution accounts for 9 ms.
- Analog-to-Digital Ingestion:** ECG pipeline acquisition adds a negligible 4 ms overhead.
- Total Worst-Case Hardware Loop:** The final local control path triggers peripheral hardware responses within 44 ms to 45 ms of an active threshold violation. Over standard cellular networks, the edge application consistently updates remote Firebase keys and delivers Twilio WhatsApp SOS alerts within 1.2 to 1.8 seconds. The Edge power consumption is sub-5W.

B. Monitored Modalities

The multi-modal monitoring system was subjected to real-time verification to assess the robustness of behavioral extraction under erratic operational postures. The posture tracking algorithm's capability to detect prolonged, atypical driver orientation was tested. During a lateral head-tilt event, which mimics micro-sleep muscle relaxation or physical incapacitation, when the relative rotation angle reaches $\theta_{\text{tilt}} = 41^\circ$, surpassing the boundary target limits ($\pm 40^\circ$), the state machine triggers a persistent alert flag within 45 milliseconds. To monitor global facial metrics simultaneously, the subsystem overlays a continuous coordinate mesh across the driver's face, tracking structural changes in the eyes and mouth zones. Even when high-priority hazards are active or state flags fluctuate, the edge node continues compiling tracking parameters (EAR = 0.29, MAR = 0.35). Furthermore, oral expansion cycles were verified to evaluate the pipeline's sensitivity to drowsiness-induced yawning patterns. When oral cavity deformation drives the Mouth Aspect Ratio upward (MAR = 0.63), breaching the strict baseline trigger limit ($\tau_M = 0.6$), the temporal counter instantly flags the anomaly status ("Yawning Alerts: 1/3"). The resilience of the ocular processing tracker was verified during a simulated

prolonged micro-sleep breakdown. When eyelid coordinates collapse due to exhaustion, the calculated Eye Aspect Ratio drops explicitly to $EAR = 0.16$. Because this ratio breaches the strict minimum safety threshold ($\tau_E = 0.2$), the temporal smoothing engine registers a validated state violation ("Sleep Alerts: 1/3"). Finally, the cognitive distraction pipeline was validated by simulating hands-on mobile multitasking behavior within the vehicle perimeter. The color segmentation matrices and adaptive masks cleanly filter skin tones away from the handheld hardware structure. This allows the system to isolate the smartphone profile and generate a high-confidence oriented bounding box hull over it (resolved at 0.86 confidence metric bounds). Because the calculated centroid of the device breaches the strict Euclidean safety parameter radius ($d_{prox} \leq 150$ pixels) measured against the driver's outer ear and jaw landmarks, the localized tracking engine immediately flags a persistent distraction event. This confirmation forces a pulse-width modulation throttle adjustment to slow the vehicle down safely.



Fig - 3: Head tilt = 41



Fig - 4: Active State



Fig - 5: Yawning Event (MAR=0.63)



Fig - 6: Blinks (EAR=0.16)

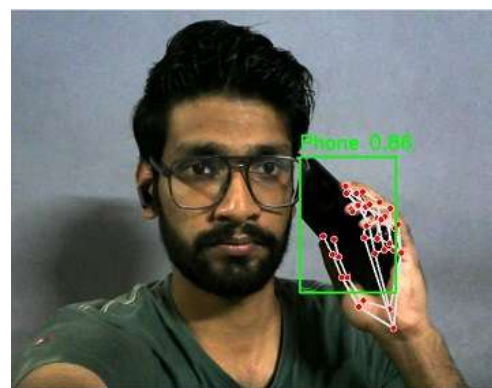


Fig - 7: Cognitive Distraction

C. Comparative Analysis with Existing Works

To quantify the multi-disciplinary contributions of this research framework, the system is cross-referenced with foundational methodologies established within the vehicular telemetry and intelligent driver assistance domains. Table 1 details these performance comparisons. The quantitative and structural synthesis of the experimental data highlights several clear technical advantages of the pro-proposed multi-modal edge intelligence framework over baseline literature benchmarks:

- True Multi-Modal Synthesis: The system uniquely unifies high-frequency behavioral computer vision tracking matrices alongside isolated physiological cardiovascular telemetry (ECG vectors). This hybrid sensing paradigm successfully bridges the single-modality vulnerabilities characteristic of traditional, isolated observation frameworks.
- Active Closed-Loop Intervention: Unlike existing architectures that rely exclusively on passive tracking arrays, diagnostic panels, or standard portable screen popups, the proposed system introduces an immediate, hardware-level automated actuation mechanism. By integrating an optically isolated pulse-width modulation (PWM) deceleration H-bridge alongside automated, out-

of-band Twilio WhatsApp SOS cellular dispatch streams, the system moves beyond basic alerts to implement real-time accident prevention.

- **Low-Latency Edge Performance:** The framework eliminates the heavy computational overhead, spatial artifact bottlenecks, and thermal throttling limits common when deploying deep learning models on standard microcontrollers or unoptimized portable smart terminals. Operating strictly within tight edge power boundaries on a Raspberry Pi 5 single-board computer, the optimized code blocks maintain a stable processing throughput of 32–35 frames per second (FPS).

- **Ultra-Fast Local Hardware Response:** By executing all primary computer vision parsing and biometric wave telemetry filtering locally on the decentralized edge node, the processing core circumvents the late-stage lag trends inherent to vehicle CAN-bus tracking methods. The total worst-case local control loop activates critical physical relays and audio-visual alarms within a mere 45 milliseconds of a validated hazard threshold violation.

- **Real-Time Sub-Second Cloud Synchronization:** Through an asynchronous, multi-threaded communication pipeline, the system resolves local cellular I/O bottlenecks. Structural JSON data payloads are streamed directly to Firebase Realtime Database clusters and out-of-band emergency alerts are executed within a 1.2 to 1.8-second threshold.

Study/Ref	Core Platform	Modalities	Cloud	Intervention	Latency/Perf
Du et al. [3]	Heavy CNN	Eyelid, PERCLOS, HR	None (Offline)	None (Passive)	Restricted by deep CNN.
Lamaazi [10]	Dist. Edge	Face tracking, maps	Passive UI	Public tel.	Constrained by network.
Delwar [14]	YOLOv5/v8	Facial Landmark	None	Local flags	Restricted view.
Pandey [26]	Mobile Host	Face Tracker	None	Alarm	Locally restricted.
Hassan [27]	Machine L.	EEG, EOG, ECG	None	None	Local detection.
Rashid [32]	Hybrid Host	EEG & EMG	Local Array	None	Prone to artifacts.
Khurana [33]	Desktop Panel	Smartphone Cam	None	Passive Log	Excessive overhead.
Our Work	Raspberry Pi 5	Ocular, ECG, Phone	Async Firebase	PWM Decel, SOS	32-35 FPS, 45ms loop.

Table -1: Comparative analysis positioning the proposed multi-modal framework against baseline literature frameworks.

VII. CONCLUSIONS

This study presents a robust, four-layered multi-modal edge intelligence framework for real-time driver telemetry monitoring and closed-loop intervention, utilizing a decentralized edge architecture to overcome the bandwidth and latency limitations of centralized systems. By pairing localized computer vision pipelines with galvanically isolated physiological electrocardiogram (ECG) telemetry modules, the framework provides a fast, dependable, and non-intrusive in-cabin safety system. Empirical validation demonstrates that the proposed architecture successfully satisfies strict automotive performance constraints. The edge computing

pipeline consistently delivers a steady 32–35 FPS spatial tracking frame rate while activating critical physical intervention mechanisms within an exceptional 44–45 ms threshold. Furthermore, out-of-band cloud synchronization and cellular emergency SOS alerts are successfully executed within sub-second thresholds without introducing processing bottlenecks to the primary local safety loops.

Future research trajectories will focus on:

- Integrating lightweight deep learning models to enhance facial landmark tracking stability under extreme low-light or night driving conditions.
- Expanding vehicle dynamics diagnostics by incorporating native On-Board Diagnostics (OBD-II) standard interface support.
- Integrating real-time GPS tracking arrays to deliver precise spatial mapping data payloads within the emergency messaging dispatch ecosystem.

ACKNOWLEDGEMENT

The authors acknowledge the Department of Electronics and Communication Engineering, Government Engineering College Wayanad, for supporting the physical infrastructure and equipment pipelines required for this research framework execution.

REFERENCES

- [1] World Health Organization. Global Status Report on Road Safety 2023. Geneva: World Health Organization; 2023.
- [2] Putilov AA, Donskaya OG. Construction of a scale for subjective assessment of momentary sleepiness. J Sleep Res. 2013;22(3):312–321.
- [3] Guanglong D, Zhang L, Kang S, Wang X, Teng S, Liu PX. A multi-modal fusion fatigue driving detection method based on heart rate and perclos. IEEE Trans Intell Transp Syst. 2022;23(11):21810–21820.
- [4] El-Nabi SA, El-Shafai W, El-Rabaie E-SM, Ramadan KF, El-Samie FEA, Mohsen S. Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review. Multimed Tools Appl. 2024; 83(3):9441–9477.
- [5] Biying F, Boutros F, Lin C-T, Damer N. A survey on drowsiness detection-modern applications and methods. IEEE Intell Transp Syst Mag. 2024;19:7279–7300.
- [6] El-Nabi SA, Ramadan KF, El-Rabaie E-SM, Emam A, El-Shafai W. A real-time design and implementation of intelligent drowsiness and fatigue recognition system for enhancing driver safety. Eng Appl Artif Intell. 2025b;162:112665.
- [7] Li N, Zhou R, Krishna B, Pradhan A, Lee H, He J, Jiang N. Non-invasive techniques for muscle

fatigue monitoring: a comprehensive survey. *ACM Comput Surv.* 2024; 56(9):1–40.

[8] M. K. Hussein, T. M. Salman, A. H. Miry, and M. A. Subhi. 'Driver drowsiness detection techniques: A survey. *Proc. 1st Babylon Int. Conf. Inf. Technol. Sci. (BICITS).* 2021;45–51.

[9] H. Iwamoto, K. Hori, K. Fujiwara, and M. Kano. Real-driving-implementable drowsy driving detection method using heart rate variability based on long short-term memory and autoencoder. *IFAC-PapersOnLine.* 2021. 54;15:526–531.

[10] Lamaazi H, Alqassab A, Fadul RA, Mizouni R. Smart edge-based driver drowsiness detection in mobile crowdsourcing. *IEEE Access* 2023;11:21863–21872.

[11] Osama F. Hassan, Ahmed F. Ibrahim, M. A. Makhoulf, B. Hafiz, Ahmed Gomaa. Intelligent driver monitoring systems: a survey of drowsiness detection technologies for road safety. *Artificial Intelligence Review.* 2026; 59(6):1–83.

[12] Ramzan M, Khan HU, Awan SM, Ismail A, Ilyas M, Mahmood A. A survey on state-of-the-art drowsiness detection techniques. *IEEE Access.* 2019;7:61904–61919.

[13] Sikander G, Anwar S. Driver fatigue detection systems: a review. *IEEE Trans Intell Transp Syst.* 2018;20(6):2339–2352.

[14] Delwar TS, Singh M, Mukhopadhyay S, Kumar A, Parashar D, Lee YM, Rahman H, Sejan MAS, Ryu JY. AI- and deep learning-powered driver drowsiness detection method using facial analysis. *Appl Sci.* 2025;15(3):1102–1125.

[15] Bamidele AA, Kamardin K, Syazarin Natasha N, et al. Non-intrusive driver drowsiness detection based on face and eye tracking. *Int J Adv Comput Sci Appl.* 2019;10(7).

[16] Turki A, Kahouli O, Albadran S, Ksantini M, Aloui A, Amara MB. A sophisticated drowsiness detection system via deep transfer learning for real time scenarios. *AIMS Math.* 2024;9(2):3211–3234.

[17] Dewi C, Chen RC, Jiang X, Yu H. Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks. *PeerJ Comput Sci.* 2022;8:e943.

[18] Mehta S, Dadhich S, Gumber S, Jadhav Bhatt A. Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio. In: *Proc. SUSCOM. Amity University Rajasthan, Jaipur-India;* 2019.

[19] Souchet AD, Philippe S, Lourdeaux D, Leroy L. Measuring visual fatigue and cognitive load via eye tracking while learning with virtual reality head-mounted displays: A review. *Int J Hum Comput Interact.* 2022;38(9):801–824.

[20] Kolus A. A systematic review on driver drowsiness detection using eye activity measures. *IEEE Access.* 2024;12:97969–97993.

[21] Al-Madani AM, Gaikwad AT, Mahale V, Ahmed ZAT, Shareef AAA. Real-time driver drowsiness detection based on eye movement and yawning using facial landmark. In: *2021 ICCI. IEEE;* 2021;1–4.

[22] Moujahid A, Dornaika F, Arganda-Carreras I, Reta J. Efficient and compact face descriptor for driver drowsiness detection. *Expert Syst Appl.* 2021;168:114334.

[23] Ansari S, Naghdy F, Haiping D, Pahnwar YN. Driver mental fatigue detection based on head posture using new modified relu-bilstm deep neural network. *IEEE Trans Intell Transp Syst.* 2021;23(8):10957–10969.

[24] Wijnands JS, Thompson J, Nice KA, Aschwanden GDPA, Stevenson M. Real-time monitoring of driver drowsiness on mobile platforms using 3d neural networks. *Neural Comput Appl.* 2020;32(13):9731–9743.

[25] Qingtian W, Li N, Zhang L, Fei Richard Yu. Driver drowsiness detection based on joint human face and facial landmark localization with cheap operations. *IEEE Trans Intell Transp Syst.* 2024;25:19633–19645.

[26] Pandey R, Bhasin P, Popli S, Sharma M, Sharma N. Driver drowsiness detection and traffic sign recognition system. In: *Emerging Technologies in Data Mining and Information Security: Proc. IEMIS 2022. Springer.* 2022; 1:25–40.

[27] Hasan MM, Watling CN, Larue GS. Validation and interpretation of a multimodal drowsiness detection system using explainable machine learning. *Comput Methods Programs Biomed.* 2024;243:107925.

[28] Stancin I, Cifrek M, Jovic A. A review of EEG signal features and their application in driver drowsiness detection systems. *Sensors.* 2021;21(11):3786.

[29] Yaacob S, Afrina Izzati Affandi N, Krishnan P, Rasyadan A, Yaakop M, Mohamed F. Drowsiness detection using EEG and ECG signals. In: *2020 IEEE 2nd IICAIET. IEEE;* 2020. pp. 1–5.

[30] Amidei A, Fallica PG, Conoci S, Pavan P. Validating photoplethysmography (ppg) data for driver drowsiness detection. In: *2021 IEEE MetroAutomotive. IEEE;* 2021. pp. 147–151.

[31] Mejía-Mejía E, Budidha K, Abay TY, May JM, Kyriacou PA. Heart rate variability (hrv) and pulse rate variability (prv) for the assessment of autonomic responses. *Front Physiol.* 2020;11:779.

[32] Rashid M, Mustafa M, Sulaiman N, Nahidul Islam Md. EEG and emg-based multimodal driver

drowsiness detection: a CWT and improved VGG-16 pipeline. In: International Human Engineering Symposium. Springer; 2023. pp. 339–350.

[33] Khurana R, Goel M. Eyes on the Road: Detecting Phone Usage by Drivers Using On-Device Cameras. In: Proc. CHI '20. New York: ACM; 2020. pp. 1–11.

[34] Perkins E, Sitaula C, Burke M, Marzbanrad F. Challenges of driver drowsiness prediction: the remaining steps to implementation. arXiv preprint arXiv:210908355. 2021.

[35] Priyanka S, Shanthi S, Saran Kumar A, Praveen V. Data fusion for driver drowsiness recognition: a multimodal perspective. Egypt Inform J. 2024;27:100529.

[36] Florez R, Palomino-Quispe F, Alvarez AB, Coaquira-Castillo RJ, Herrera-Levano JC. A real-time embedded system for driver drowsiness detection based on visual analysis of the eyes and mouth using convolutional neural network and mouth aspect ratio. Sensors. 2024;24(19):6261.