iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

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

DOI: https://doi.org/10.22214/ijraset.2026.79348

Real-Time Driver Distraction Detection System Using Embedded Vision

Venkat Raman T[1], Mollish Kreson R[2], Syed Roshan Ali R[3], Noorul Jameel H[4], Vasu G[5]
[1, 2, 3, 4]UG-Scholar, [5]Assistant Professor, Information Technology, M.I.E.T. Engineering College, Tiruchirappalli

Abstract: Driver inattention and fatigue significantly contribute to road accidents, especially during prolonged and high-speed driving conditions. This work presents an embedded real-time driver distraction detection system built using the ESP32-S3 CAM module. The system evaluates driver behavior by monitoring eye closure duration and head movement patterns using computationally efficient techniques suitable for resource-constrained hardware.

A key aspect of the proposed system is the integration of vehicle speed data obtained from a GPS module to dynamically adjust detection thresholds. Additionally, an event-based alert mechanism is introduced to identify repeated distraction occurrences, combined with a cooldown strategy to avoid excessive alerts and improve user experience.

The system operates as a self-contained embedded unit without reliance on external processing or network connectivity. Experimental observations indicate stable performance with reduced false alerts, demonstrating its applicability for real-time driver safety monitoring.

Keywords: Driver Monitoring, Embedded Vision, ESP32-S3 CAM, Distraction Detection, Adaptive Alert System, Real-Time Processing

I. INTRODUCTION

Driver inattention remains a critical factor contributing to accident risk in modern transportation systems of road accidents worldwide, contributing significantly to fatalities and injuries in both urban and highway environments. According to recent transportation safety studies, a considerable percentage of accidents occur due to driver fatigue, inattention, and delayed reaction time caused by distraction-related behaviors [1]. Distraction may arise from various factors including prolonged eye closure, frequent head movement away from the road, and reduced situational awareness. These factors directly affect the cognitive and visual attention of the driver, leading to critical safety risks.

Traditional vehicle safety systems mainly focus on mechanical and environmental factors such as braking systems, lane detection, and collision avoidance mechanisms. However, these systems fail to monitor the physiological and behavioral state of the driver, which is equally important for ensuring road safety [2].

Advanced driver monitoring systems available in modern vehicles often rely on high-end hardware and computationally intensive algorithms, making them unsuitable for low-cost and embedded applications. In recent years, computer vision-based approaches have gained attention for real-time driver monitoring due to their non-intrusive nature and ability to analyze facial features [3]. Techniques such as eye state detection and head pose estimation have been widely explored for identifying signs of driver fatigue and distraction. However, many of these approaches depend on deep learning models that require high computational resources and are not feasible for embedded systems with limited processing capabilities. To overcome these limitations, this paper proposes a real-time driver distraction detection system using embedded vision implemented on the ESP32-S3 CAM module. The system utilizes lightweight detection techniques to monitor eye closure duration and head movement patterns without relying on heavy machine learning models. This approach ensures efficient real-time performance within the constraints of embedded hardware. Furthermore, the proposed system integrates a GPS module to obtain real-time vehicle speed, enabling adaptive threshold-based decision making. The alert mechanism is designed to provide immediate feedback through a buzzer and a voice alert system, ensuring that the driver is promptly notified during distraction events. An event-based alert strategy is implemented to avoid unnecessary alerts and improve system usability. The key objective of this work is to develop a cost-effective, standalone, and efficient driver monitoring system that can operate without external processing units or internet connectivity. By combining lightweight embedded vision techniques with intelligent alert mechanisms, the proposed system aims to enhance road safety and reduce accident risks.

The novelty of the proposed system lies in combining lightweight embedded vision techniques with adaptive speed-based thresholds and event-driven alert mechanisms, enabling efficient and practical real-time driver monitoring.

II. LITERATURE SURVEY

Driver distraction detection has been an active area of research in the field of intelligent transportation systems and computer vision. Various techniques have been proposed to monitor driver behavior using image processing, machine learning, and sensor-based approaches. Early research focused on rule-based and geometric methods for detecting driver fatigue by analyzing eye blinking patterns and facial landmarks. These approaches relied on predefined thresholds and handcrafted features, which limited their robustness under varying lighting and environmental conditions [4].

Although computationally efficient, these methods lacked adaptability and accuracy in real-world scenarios. With the advancement of machine learning, researchers introduced classification-based approaches for driver monitoring. Techniques such as Support Vector Machines (SVM) and Random Forest classifiers were used to analyze facial features and detect driver states [5].

These methods improved detection accuracy but required feature extraction pipelines and labeled datasets, increasing system complexity. Deep learning-based approaches have further enhanced performance by enabling automatic feature extraction using Convolutional Neural Networks (CNNs). Several studies have demonstrated high accuracy in detecting eye state and head pose using deep learning models trained on large datasets [6].

However, these models require significant computational power, making them unsuitable for deployment on embedded systems such as microcontrollers. Recent research has focused on embedded vision systems that aim to balance performance and computational efficiency. Lightweight algorithms and optimized processing techniques have been proposed to enable real-time operation on low-power devices [7].

However, many existing systems lack adaptive mechanisms to adjust detection thresholds based on driving conditions. Additionally, most systems generate alerts for every detected event, leading to excessive notifications and reduced user acceptance. Studies have shown that repeated alerts can cause driver annoyance and reduce system effectiveness [8]. Therefore, there is a need for intelligent alert strategies that consider event frequency and context. The proposed system addresses these gaps by integrating lightweight detection methods with adaptive speed-based logic, event-based alert mechanisms, and cooldown strategies. This combination provides a practical and efficient solution for real-time driver distraction detection in embedded environments.

III.METHODOLOGY

The proposed driver distraction detection system follows a structured methodology consisting of image acquisition, feature analysis, decision making, and alert generation. The system is designed to operate entirely on the ESP32-S3 CAM module, ensuring a fully embedded implementation without external computational support. The process begins with continuous image acquisition using the onboard camera module. The captured frames are processed in real time to extract relevant features associated with driver behavior. The system focuses on two primary indicators: eye closure duration and head movement patterns. Eye state detection is performed using a threshold-based approach, where the system analyzes the visibility of the eye region across consecutive frames. If the eyes remain closed beyond a predefined duration, the system identifies it as a potential distraction event. This approach eliminates the need for complex machine learning models while maintaining reliable performance.

Head movement detection is implemented by analyzing positional variations of the face across frames. Significant deviation from the forward-facing position indicates that the driver is not focusing on the road. The system continuously tracks these movements and evaluates their duration to determine distraction.

To enhance system reliability, filtering techniques are applied to remove noise and avoid false detections. Short-duration events are ignored, and only sustained behaviors are considered for further processing. This ensures stable operation under real-world conditions. A key feature of the system is the integration of speed-based adaptive logic. The GPS module provides real-time speed data, which is used to adjust detection thresholds dynamically.

At lower speeds, the system reduces sensitivity to avoid unnecessary alerts, while at higher speeds, stricter thresholds are applied to ensure timely warnings. The system employs an event-based detection strategy to analyze repeated distraction occurrences. Instead of triggering alerts for every detection, the system monitors the frequency of events within a defined time window. If multiple events occur within this period, the system escalates the alert level. The alert mechanism consists of two stages. Initially, a buzzer alert is triggered to provide immediate feedback.

If repeated events are detected, a voice alert is generated using a DFPlayer Mini module. A cooldown period is implemented after voice alerts to prevent excessive notifications and improve user experience. This multi-stage methodology ensures accurate detection, efficient processing, and controlled alert generation, making the system suitable for real-time embedded applications.

A. System Flowchart

Fig. 3 illustrates the operational flow of the proposed driver distraction detection system. The system begins by initializing all hardware components, including the ESP32-S3 CAM module and GPS module. Once initialized, the system continuously captures image frames of the driver and processes them to detect eye closure and head movement. The detection results are passed through a filtering stage to eliminate noise and ensure stable operation.

The system then reads real-time speed data from the GPS module and applies speed-based threshold logic to determine whether the detected behavior qualifies as distraction. If a distraction condition is identified, the system triggers a buzzer alert for immediate feedback. The system also maintains a count of distraction events within a predefined time window. If the number of events exceeds the defined threshold, a voice alert is generated to notify the driver. After a voice alert is triggered, a cooldown period is applied to prevent repeated alerts. The system then returns to continuous monitoring mode, ensuring real-time operation.

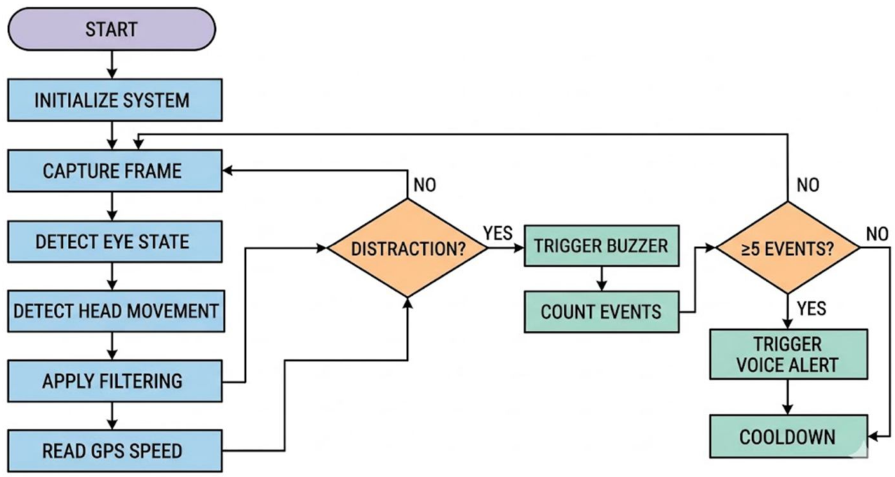


Fig. 1 System Flowchart

IV. SYSTEM ARCHITECTURE

The proposed system architecture is designed as a standalone embedded driver monitoring system that integrates vision processing, decision logic, and alert mechanisms within a compact hardware framework. Unlike conventional systems that rely on external processing units, the entire computation is performed onboard using the ESP32-S3 CAM module. The architecture consists of four major layers: image acquisition layer, processing layer, decision layer, and alert generation layer.

These layers work sequentially to ensure efficient real-time operation under constrained computational resources. The image acquisition layer is responsible for capturing continuous video frames using the camera integrated with the ESP32-S3 CAM module. The captured frames are processed at a controlled frame rate to balance computational load and detection accuracy. Since embedded systems have limited processing power, frame sampling techniques are employed to avoid unnecessary processing of redundant frames. The processing layer performs feature extraction and behavioral analysis.

In this system, lightweight techniques are used to detect eye state and head orientation. Instead of relying on deep learning models, threshold-based detection methods are implemented to ensure faster execution and reduced memory consumption.

|This design choice significantly improves real-time performance and makes the system suitable for embedded deployment [1]. The decision layer evaluates the extracted features based on predefined conditions. The system continuously monitors eye closure duration and head deviation patterns.

These parameters are compared against adaptive thresholds that vary depending on the vehicle speed obtained from the GPS module. The decision-making process ensures that alerts are triggered only when the distraction persists beyond acceptable limits. The alert generation layer consists of a buzzer and a voice alert module. The buzzer provides immediate feedback for short-term distraction events, while the voice alert system is activated for prolonged or repeated distraction occurrences. The DFPlayer Mini module is used to play pre-recorded voice alerts, ensuring clear communication with the driver.

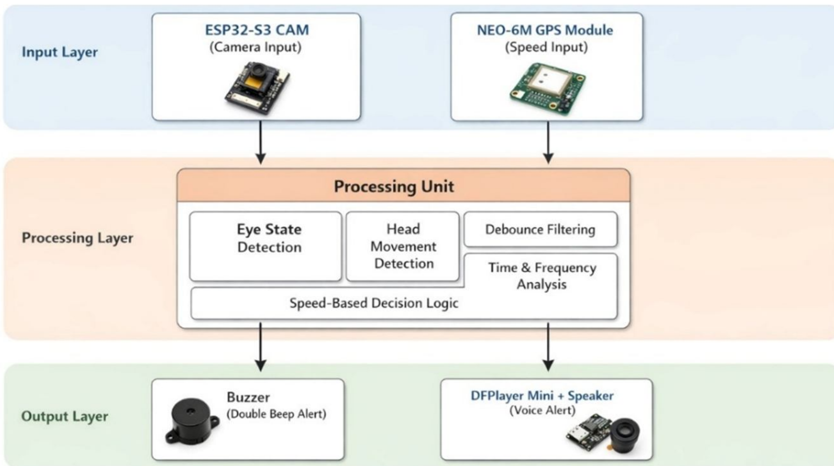


Fig. 2 System Architecture

As shown in Fig. 2, the proposed system is structured into input, processing, and output layers, enabling efficient real-time monitoring and alert generation. The system architecture shown in Figure 2 represents a layered embedded framework designed for real-time driver distraction detection.

The architecture is divided into three primary layers: input layer, processing layer, and output layer. Each layer performs a specific function and contributes to the overall operation of the system. The input layer consists of the ESP32-S3 CAM module and the NEO-6M GPS module. The ESP32-S3 CAM serves as the primary sensing unit, capturing real-time image frames of the driver's face.

These frames are used to analyze driver behavior continuously. The GPS module provides real-time speed data, which is a critical parameter for adaptive decision-making. The combination of visual input and speed data enables the system to operate in a context-aware manner.

The processing layer is the core of the system and is responsible for analyzing the captured data. It includes multiple functional blocks such as eye state detection, head movement detection, debounce filtering, time and frequency analysis, and speed-based decision logic. Eye state detection identifies whether the driver's eyes are open or closed, while head movement detection monitors deviations from the forward-facing position.

These two parameters serve as primary indicators of driver distraction. To ensure stability and reliability, debounce filtering is applied to eliminate noise and prevent false detections caused by rapid fluctuations in input data. The time and frequency analysis module evaluates the duration and repetition of distraction events. This allows the system to distinguish between occasional behavior and consistent unsafe driving patterns.

The speed-based decision logic integrates the input from the GPS module and dynamically adjusts detection thresholds. This ensures that the system is less sensitive at low speeds and more responsive at higher speeds, improving overall efficiency and reducing unnecessary alerts. The output layer consists of a buzzer and a DFPlayer Mini module connected to a speaker. The buzzer provides immediate alerts in the form of a double beep when distraction is detected.

The DFPlayer Mini module is used to generate voice alerts for repeated or prolonged distraction events. This multi-level alert system enhances driver awareness and improves safety. The layered architecture ensures modularity, scalability, and efficient real-time performance. By integrating sensing, processing, and alert generation within a single embedded system, the proposed architecture eliminates the need for external processing units and provides a compact and cost-effective solution.

V. SYSTEM DESIGN AND LOGIC

The system design focuses on developing an efficient and reliable logic for detecting driver distraction based on behavioral patterns. The detection logic is structured into multiple stages to ensure accuracy and avoid false positives. The first stage involves continuous monitoring of eye state. The system tracks whether the driver's eyes are open or closed over consecutive frames. If the eyes remain closed for a duration exceeding a predefined threshold, the system identifies it as a distraction event. This threshold is dynamically adjusted based on vehicle speed to ensure context-aware detection.

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

The second stage involves head movement analysis. The system evaluates the orientation of the driver's head and detects deviations from the forward-facing position. If the driver frequently looks away from the road or maintains a non-forward position for a prolonged duration, it is classified as a distraction. To improve reliability, the system incorporates debounce filtering techniques. This ensures that short-term fluctuations or noise in detection do not trigger false alerts.

Only consistent and sustained patterns are considered valid distraction events A key feature of the system is the implementation of speed-based adaptive thresholds. The GPS module provides real-time speed data, which is used to adjust detection sensitivity. At lower speeds, the system allows greater tolerance to avoid unnecessary alerts. As the speed increases, the system becomes more sensitive, reducing the allowable distraction duration.

The system follows the threshold logic defined as:

- ≤10 km/h → No alert
- 10–40 km/h → 5 seconds
- 40–60 km/h → 4 seconds
- 60–80 km/h → 3 seconds
- 80 km/h → 2 seconds

This adaptive mechanism ensures that alerts are proportional to driving risk levels [2]. Additionally, the system implements an event-based logic for triggering voice alerts. If distraction events occur repeatedly within a short time window (e.g., 5 times within 30 seconds), the system escalates the alert level and activates voice feedback.

To prevent excessive alerts, a cooldown mechanism is introduced. After a voice alert is triggered, the system temporarily disables further voice alerts for 60 seconds. This prevents driver annoyance and improves system usability. The combination of adaptive thresholds, event-based logic, and cooldown strategy ensures a balanced and effective detection system.

A. Adaptive Threshold and Alert Strategy

The system implements an adaptive threshold mechanism that dynamically adjusts detection sensitivity based on vehicle speed. This approach ensures that alerts are generated only when necessary and are aligned with driving risk levels.

At low speeds (≤10 km/h), the system does not trigger alerts, as the risk associated with distraction is minimal. As the speed increases, the allowable duration for distraction decreases. For speeds between 10–40 km/h, a threshold of 5 seconds is used, while for speeds above 80 km/h, the threshold is reduced to 2 seconds.

This ensures faster response in high-risk conditions. In addition to duration-based detection, the system incorporates an event-based strategy. The system maintains a count of distraction events within a predefined time window of 30 seconds. If the number of events exceeds five within this period, the system escalates the alert level and triggers a voice alert.

To prevent excessive alerting, a cooldown mechanism is implemented. After a voice alert is triggered, the system disables further voice alerts for a duration of 60 seconds. This prevents repeated notifications and improves driver comfort. This combination of adaptive thresholds, event-based analysis, and cooldown control significantly enhances system reliability and usability.

VI. SYSTEM IMPLEMENTATION

The implementation of the proposed system is carried out using embedded hardware components and optimized software techniques. The ESP32-S3 CAM module serves as the core processing unit, handling both image acquisition and real-time analysis. The camera module continuously captures frames, which are processed using lightweight algorithms. The processing is optimized to minimize latency and ensure real-time performance. Efficient memory management techniques are applied to handle image data within the limited RAM available on the microcontroller. The GPS module (NEO-6M) is interfaced with the ESP32 to obtain real-time speed data. The speed information is parsed and used to dynamically adjust detection thresholds. This integration enables context-aware decision making. The alert system is implemented using two components: an active buzzer and a DFPlayer Mini module connected to a speaker. The buzzer is triggered using GPIO signals, while the DFPlayer Mini plays pre-recorded audio messages stored on an SD card. The use of pre-recorded audio ensures clarity and reduces processing overhead. The system also incorporates a buck converter to provide stable voltage levels to all components. Proper voltage regulation is essential to prevent system instability and ensure reliable operation. The software implementation is designed to run efficiently on the ESP32 platform. The code is structured into modular functions for image processing, decision making, and alert control. Interrupt-based and timer-based mechanisms are used to manage real-time operations. Overall, the implementation demonstrates that complex driver monitoring functionalities can be achieved using lightweight embedded systems without the need for high-end hardware.

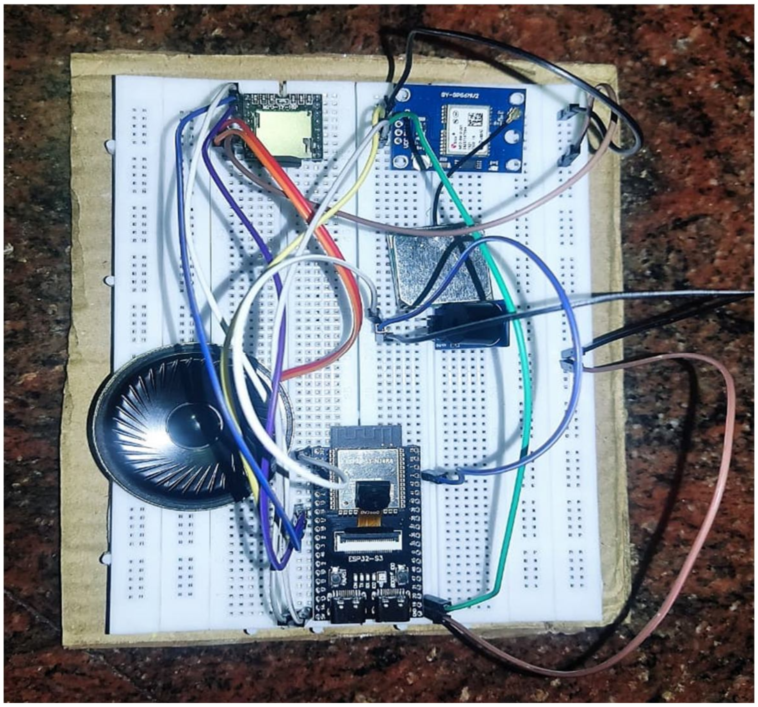


Figure 2: Physical Prototype of the Embedded Driver Distraction Detection System

The hardware implementation of the system is depicted in Fig. 3, which shows the integration of all modules on a compact breadboard setup. The physical prototype of the proposed system is shown in Figure 3. The hardware setup is implemented on a breadboard and consists of the ESP32-S3 CAM module, NEO-6M GPS module, DFPlayer Mini audio module, speaker, and buzzer interconnected using jumper wires.

The ESP32-S3 CAM module acts as the central processing unit and is responsible for capturing image frames and executing the detection algorithms. The GPS module is connected to the ESP32 to provide real-time speed data, which is used for adaptive threshold adjustment. The DFPlayer Mini module is interfaced with the ESP32 to generate voice alerts through the connected speaker, while the buzzer provides immediate audible warnings.

The wiring connections are arranged to ensure proper communication between modules using serial interfaces and GPIO pins. A regulated power supply is provided to maintain stable operation of all components. The compact arrangement of components demonstrates that the system can be implemented using low-cost hardware with minimal space requirements.

The prototype validates the feasibility of the proposed system and confirms that real-time driver monitoring can be achieved using embedded vision techniques without the need for external processing units.

A. Hardware Integration Details

The prototype illustrates the integration of multiple hardware components into a single embedded system. The ESP32-S3 CAM module is interfaced with peripheral modules using UART and GPIO communication. The GPS module communicates speed data through serial communication, while the DFPlayer Mini uses serial commands to control audio playback.

The speaker is connected to the DFPlayer Mini to output voice alerts, and the buzzer is directly controlled by the ESP32 for immediate feedback. The use of a breadboard allows flexible prototyping and easy modification of connections during development. Proper wiring and power management are essential to ensure reliable system performance. The use of a buck converter helps maintain consistent voltage levels, preventing fluctuations that may affect system operation.

VII. RESULTS AND DISCUSSION

The proposed system was tested under various conditions to evaluate its performance in detecting driver distraction. The system demonstrated reliable operation in real-time scenarios, accurately identifying both eye closure and head movement-based distractions. The eye state detection mechanism effectively identified prolonged eye closure events, which are commonly associated with driver drowsiness.

The system showed consistent performance under normal lighting conditions and frontal face orientation. Head movement detection also produced satisfactory results, accurately detecting deviations from the forward-facing position. The combination of eye and head monitoring improved overall detection reliability. The integration of speed-based adaptive thresholds significantly enhanced system performance. At higher speeds, the system responded faster to distraction events, ensuring timely alerts. At lower speeds, unnecessary alerts were minimized, improving user experience.

The buzzer alert provided immediate feedback, while the voice alert system effectively handled repeated distraction events. The cooldown mechanism successfully prevented excessive alerts, ensuring that the system remained user-friendly. However, certain limitations were observed during testing.

The system performance was affected under low-light conditions and extreme head orientations. Additionally, rapid lighting changes and occlusions sometimes reduced detection accuracy. Despite these limitations, the system achieved a high level of reliability and demonstrated the feasibility of implementing real-time driver monitoring using embedded vision techniques.

A. System Performance Evaluation

The performance of the proposed system is evaluated based on detection accuracy, response time, and reliability. The system demonstrates consistent performance in detecting both eye closure and head movement-based distractions. The response time of the system is significantly improved due to the use of lightweight processing techniques.

The adaptive threshold mechanism ensures that alerts are generated promptly at higher speeds while minimizing unnecessary alerts at lower speeds. The implementation of debounce filtering reduces false positives and improves stability. The event-based alert mechanism further enhances system performance by focusing on repeated distraction patterns rather than isolated events. Overall, the system achieves a balance between computational efficiency and detection accuracy, making it suitable for real-time embedded applications.

B. Comparative Discussion

The proposed system provides a lightweight alternative to deep learning-based driver monitoring systems. While deep learning approaches offer high accuracy, they require significant computational resources. In contrast, the proposed system achieves efficient real-time performance using embedded hardware with minimal resource requirements.

The adaptive threshold mechanism and event-based alert strategy further improve system usability by reducing false alerts and ensuring context-aware operation. This makes the system suitable for practical deployment in real-world conditions.

VIII. CONCLUSION

This paper presented a real-time driver distraction detection system using embedded vision implemented on the ESP32-S3 CAM module. The system successfully integrates lightweight detection techniques, adaptive logic, and intelligent alert mechanisms to provide an efficient and practical solution for driver monitoring.

Unlike conventional systems that rely on computationally intensive models, the proposed approach utilizes threshold-based methods to achieve real-time performance on embedded hardware. The integration of GPS-based speed adaptation enhances detection accuracy and ensures context-aware operation.

The multi-stage alert system, consisting of buzzer and voice feedback, effectively notifies the driver while maintaining usability through cooldown mechanisms. The system demonstrates that reliable driver monitoring can be achieved using low-cost hardware and optimized algorithms.

Future work can focus on improving detection accuracy under challenging conditions such as low lighting and occlusions. The integration of additional sensors and advanced algorithms can further enhance system performance and expand its applicability.

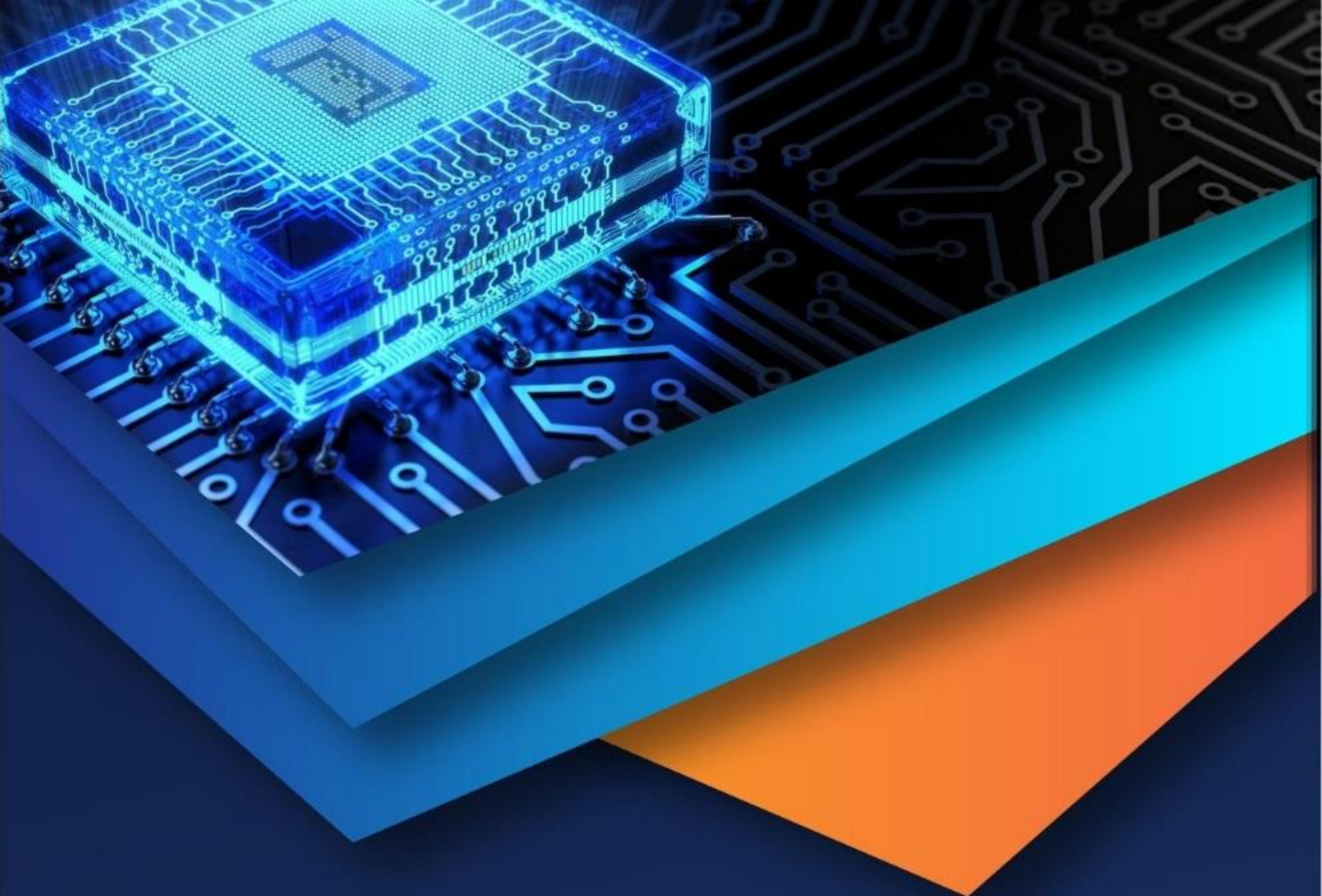
IX. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the management and faculty members of M.I.E.T Engineering College for providing the necessary support and guidance throughout the development of this project. The authors also thank the Department of Electronics and Communication Engineering for providing laboratory facilities and technical assistance.

Special appreciation is extended to the project guide for valuable suggestions and continuous encouragement, which helped in the successful completion of this work.

REFERENCES

[1] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," NHTSA, 2015.
[2] A. Doshi and M. M. Trivedi, "On the roles of eye gaze and head dynamics in predicting driver's intent," IEEE Transactions, 2009.
[3] M. Abtahi et al., "Driver drowsiness monitoring based on yawning detection," IEEE, 2011.
[4] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," CVWW, 2016.
[5] S. Kaplan et al., "Driver behavior analysis for distraction detection," Elsevier, 2015.
[6] Y. LeCun et al., "Deep learning," Nature, 2015.
[7] K. Zhang et al., "Joint face detection and alignment using MTCNN," IEEE, 2016.
[8] R. Grace et al., "A drowsy driver detection system for heavy vehicles," IEEE, 2001.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY