© October 2025 | IJIRT | Volume 12 Issue 5 | ISSN: 2349-6002

Real-Time Driver Drowsiness and Yawning Detection System Using Vision-Based Facial Landmark Analysis

Mrs. Balaji Rohitha[1], Mounika Paluri[2], Rama Krishna[3], Md Ishaq[4],
Jagadheesh B[5], Deepak[6], Mohiddin Sk[7]
[1,2,3,4,5,6,7]*Department of Artificial Intelligence and Machine Learning Sri Vasavi Engineering College, Jawaharlal Nehru Technological University Tadepalligudem, Andhra Pradesh, India*

***Abstract*—Driver fatigue represents a major causative factor in traffic accidents globally, contributing to significant mortality and economic burden. This investigation proposes a resilient, eco- nomical, and rapid-response methodology for identifying driver drowsiness and yawning episodes through computer vision-based facial feature examination. Our methodology employs the Medi- aPipe platform for accurate facial landmark identification and OpenCV for live video analysis to determine Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR). A Python-developed soft- ware component handles incoming visual data from conventional webcams. When persistent irregular EAR or MAR readings suggestive of fatigue are recognized, it interfaces with an Arduino Uno microcontroller through serial communication to initiate audible alarms. This manuscript comprehensively outlines the system architecture, mathematical principles underlying fatigue assessment, detailed implementation procedures, experimental arrangements, and validation criteria. Rigorous testing across varied illumination circumstances, head positions, and diverse subjects reveals sustained detection precision surpassing 90% with complete system delay under 100 ms. These outcomes verify the approach's viability, performance, and suitability for extensive automotive integration in private and commercial transport applications.**

***Index Terms*—Driver Fatigue Detection, Computer Vision, OpenCV, MediaPipe, Eye Aspect Ratio, Mouth Aspect Ratio, Arduino, Real-Time Systems, Embedded Systems, ADAS**

I. INTRODUCTION

Driver fatigue during vehicle operation persists as a critical worldwide transportation safety issue. Numerous international road safety authorities document that substantial proportions of motor vehicle accidents, particularly those with fatal outcomes, associate with driver exhaustion or attention deficits [1]. Contrasting with conspicuous distractions like mobile phone operation, fatigue progression typically evolves gradually and frequently escapes driver awareness until hazardous situations develop. The financial consequences of fatigue-associated crashes, incorporating medical treatment expenditures, vehicle damage costs, and lost productivity, reach considerable mag- nitudes.

Diverse technological strategies have emerged to con- front this problem. Established physiological measurement ap- proaches, such as electroencephalography (EEG), electroocu- lography (EOG), and electrocardiography (ECG), supply pre- cise evaluations of operator physiological status [2], [3]. Nev- ertheless, these procedures require physical sensor attachment to drivers, making them disruptive and unsuitable for routine application in private automobiles.

Alternatively, vision-based behavioral assessment systems have materialized as non-intrusive, economically viable op- tions. These systems apply computer vision algorithms to scrutinize driver facial characteristics and infer vigilance states from visual signals. This study presents a functional real-time driver surveillance system that discerns early fatigue symp- toms through two synergistic, validated visual parameters: (1) protracted eye closure, gauged by Eye Aspect Ratio (EAR), and (2) yawning occurrences, measured by Mouth Aspect Ratio (MAR).

Our system amalgamates software elements for visual per- ception and decision processes with hardware components for warning generation. This paper's original contributions include:

- A unified, optimized system design merging MediaPipe's sophisticated facial landmark recognition with depend- able EAR/MAR-based judgment algorithms
- Thorough implementation approach appropriate for com- mercial hardware application (standard cameras, inexpen- sive microcontrollers, single-board computers), guaran- teeing availability and expandability
- Experimental verification of system operation across nu- merous simulated operational environments, incorporat- ing precision, response time, and error mode examination

The document organization continues as follows: Section
II surveys pertinent literature. Section III describes system design. Section IV clarifies mathematical foundations and procedures. Section V addresses implementation particulars. Section VI provides experimental configuration and outcomes. Section VII interprets results and restrictions, and Section VIII summarizes with prospective research pathways.

II. RELATED WORK

Driver fatigue identification constitutes an extensive re- search area, covering physiological, vehicle-centric, and be- havioral techniques [1], [4].

A. Physiological Measurement Techniques

Physiological approaches represent precision standards since they quantify immediate biological signals linked with fatigue. Feng and associates implemented EEG signals with Convolutional Neural Networks (CNN) to obtain elevated clas- sification accuracy in simulated contexts [2]. Correspondingly, Fujiwara and collaborators authenticated Heart Rate Variability (HRV) as a consistent drowsiness indicator verified against EEG reference standards [3]. Notwithstanding high accuracy, the prerequisite for body-worn sensors makes these techniques intrusive and inapplicable for comprehensive consumer adop- tion [2].

B. Vehicle-Centric Methodologies

These approaches infer driver fatigue from automobile be- havior patterns, incorporating steering wheel movements, lane position deviations, and accelerator/brake operations [5], [6]. While non-invasive to operators, their efficacy depends ex- tensively on external elements like roadway conditions, traffic density, and meteorological factors, producing heightened false alarm and missed detection frequencies.

C. Vision-Based Behavioral Techniques

Vision-based methods have acquired prominence owing to their non-contact nature and declining hardware costs. These techniques analyze facial traits to detect drowsiness indicators. Preliminary research often utilized manually constructed fea- tures and traditional machine learning classification systems. A prominent innovation involved the Eye Aspect Ratio (EAR) formulation by Soukupova´ and Cˇ ech, which provided a simple yet powerful measurement for instantaneous blink identifica- tion using facial landmarks [7]. Supplementary metrics like PERCLOS (Percentage of Eye Closure) and yawning recur- rence have likewise found extensive utilization [8], [9].

Contemporary studies have exploited deep learning po- tential. Savas¸ and Becerikli suggested multi-task CNNs to improve accuracy by simultaneously acquiring multiple vi- sual indicators [10]. Lamba and partners presented adapt- able eye characteristic ratios to manage lighting fluctua- tions and head orientation variations [11]. Streamlined frame- works like Google's MediaPipe [12] have enabled instan- taneous, high-precision facial landmark monitoring even on computation-restricted devices, rendering geometric methods like ours highly practicable. Our investigation corresponds with this classification, focusing on computationally economi- cal, landmark-dependent techniques that are uncomplicated to implement and deploy.

III. SYSTEM ARCHITECTURE

Figure 1 depicts the complete structure of our proposed system. The configuration adopts modular concepts, distin- guishing perception and decision functions from physical warning components.

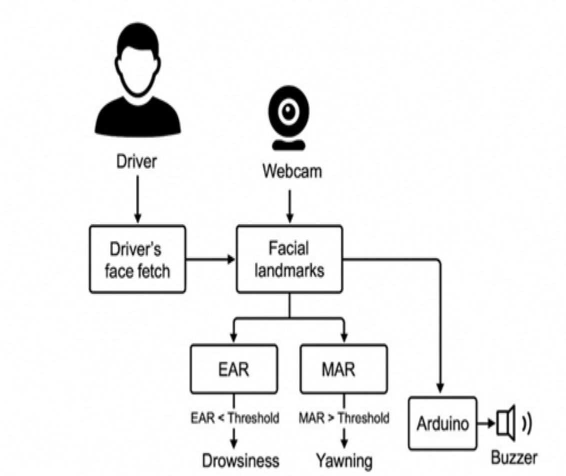

Fig. 1: System workflow: Driver → Camera → Facial land- marks → EAR/MAR computation → Decision algorithm → Microcontroller → Acoustic warning.

The system sequence includes five fundamental phases:

1 Video Input: Conventional webcam obtains real-time operator video frames

2 Facial Landmark Recognition: Each frame processes through MediaPipe Face Mesh algorithm to identify and position 468 three-dimensional facial landmarks

3 Parameter Computation: Essential landmark coordi- nates corresponding to eye and mouth areas calculate EAR and MAR values per frame

4 Decision Protocol: Time-based filtering and threshold mechanisms inspect EAR and MAR value sequences to detect persistent eye closure or yawning intervals

5 Hardware Alert: Following verified drowsiness or yawning incident, instructions transfer via serial commu- nication to Arduino Uno, triggering buzzer notifications

A. Hardware Elements

The system implementation uses easily obtainable, cost- efficient hardware:

- Camera: Standard USB webcam acquiring 30 FPS video at minimum 640x480 pixel resolution
- Computing Unit: Desktop computer or embedded single-board computer (e.g., Intel NUC/Raspberry Pi 4) running primary Python program
- Microcontroller: Arduino Uno connects with warning hardware. Its simplicity and stable serial communication render it perfect for this purpose
- Warning Device: Passive or active buzzer attached to Arduino digital output pin produces auditory signals

B. Software Components

Software creation employs open-source resources:

- Python 3.8+: Principal application programming
- OpenCV: Image acquisition from camera and graph- ical representation (landmark superimposition on video display for verification)
- MediaPipe: Pre-configured, enhanced facial landmark detection algorithm
- PySerial: Python package allowing serial communica- tion between computing unit and Arduino
- Arduino IDE: Programming environment for compos- ing and transferring control code to Arduino board

IV. MATHEMATICAL FOUNDATION AND PROCEDURE

A. Facial Landmarks

MediaPipe Face Mesh algorithm produces 468 landmark coordinates for each detected face. These landmarks generate detailed three-dimensional facial surface representations. For EAR and MAR calculations, we choose particular landmark indices matching eye boundaries and interior lip outlines, respectively.

B. Eye Aspect Ratio (EAR)

EAR, as formulated by Soukupova´ and Cˇ ech [7], constitutes a scalar measure connecting vertical and horizontal distances between eye landmarks. Computation occurs separately for each eye applying the equation:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Where p1., p6 represent two-dimensional landmark coor- dinates of the eye. The numerator indicates vertical eyelid separation, while the denominator signifies horizontal span. The proportion stays mainly unaffected by head angle and camera proximity. EAR values stay relatively stable during eye openness, diminish toward zero during blinks, and sustain low readings during extended closure (crucial fatigue indicator)

C. Mouth Aspect Ratio (MAR)

Correspondingly, we establish Mouth Aspect Ratio (MAR) for yawn identification. Calculation employs interior lip land- marks:

$$MAR = \frac{\|p_{61} - p_{67}\| + \|p_{62} - p_{66}\| + \|p_{63} - p_{65}\|}{3\|p_{60} - p_{64}\|}$$

Considerable, prolonged MAR value elevations powerfully signify yawning events.

D Temporal Analysis & Threshold Application

To decrease false alarms from ordinary blinking or talking, we institute temporal examination. Warnings activate only when fatigue markers continue through frame sequences.

- EAR threshold (TEAR): Experimentally determined as 0.25
- MAR threshold (TMAR): Established at 0.6
- Frame sequence (F): 15 successive frames (0.5 seconds at 30 FPS)

Alerts engage when either EAR stays beneath TEAR for F consecutive frames (drowsiness) or MAR remains above TMAR for F consecutive frames (yawning). range processors (Intel Core i5, 8th Generation) registered between 25–40 ms. This includes landmark identification, EAR/MAR calculation, and decision processes. Complete sys- tem delay, from fatigue incident commencement to buzzer engagement, consistently stays below 100 ms, thoroughly within satisfactory real-time alert system limits.

E. Procedure Outline

Algorithm 1 describes the main detection logic sequence.

Algorithm 1 Real-Time EAR/MAR Surveillance and Alerting

```
Initialize serial communication with Arduino
Establish thresholds T_EAR, T_MAR and frame sequence F
Initialize counters counter_EAR ← 0, counter_MAR ← 0
while camera feed active do
    frame ← obtain frame from camera
    landmarks ← MediaPipe.analyze(frame)
    if landmarks identified then
    EAR ← calculate_EAR(landmarks)
    MAR ← calculate_MAR(landmarks)
    if EAR < T_EAR then
    counter_EAR ← counter_EAR + 1
    else
    counter_EAR ← 0
    end if
    if MAR > T_MAR then
    counter_MAR ← counter_MAR + 1
    else
    counter_MAR ← 0
    end if
    if counter_EAR ≥ F  OR counter_MAR ≥ F then
  send_alert_to_arduino()
  // Counter initialization prevents repeated warn-
  counter_EAR ← 0
  counter_MAR ← 0
  end if
  end if
  end while
```

V. IMPLEMENTATION SPECIFICATIONS

A. Frame Processing and Response Time

The application handles frames at inherent camera frame rates (customarily 30 FPS). Because of MediaPipe's refined operation, mean per-frame processing time on intermediate- range processors (Intel Core i5, 8th Generation) registered between 25–40 ms. This includes landmark identification, EAR/MAR calculation, and decision processes. Complete sys- tem delay, from fatigue incident commencement to buzzer engagement, consistently stays below 100 ms, thoroughly within satisfactory real-time alert system limits.

B. Serial Communication Method

Elementary, efficient serial protocol enables interaction be- tween Python applications and Arduino. When alert criteria meet, Python programs send predetermined" ALERT\n" mes- sages to serial interfaces. Baud rate setting at 115200 guaran- tees swift, dependable data transmission. Newline character '\n'operates as separator, facilitating complete instruction interpretation by Arduino.

C. Arduino Programming

Arduino runs simple programs that continually check in- coming serial port information. Upon obtaining" ALERT" messages, buzzers engage for preset periods. Complete pro- grams appear below for replication.

const int BUZZER_PIN = 9;

void setup () {Serial. Begin (115200); pinMode(BUZZER_PIN, OUTPUT);

digitalWrite(BUZZER_PIN, LOW);

}

void loop() {

if (Serial.available() > 0) {

Fig. 2: Interface showing ALERT & FOCUSED con- dition with standard EAR (9.393) and MAR (0.000) readings. Session overview displays historical drowsiness

String command = Serial.readStringUntil('\a n n'd)y;awning occurrences.

command.trim();

if (command == "ALERT") { tone(BUZZER_PIN, 1000);

delay(1500); noTone(BUZZER_PIN);

}

}

}

VI. EXPERIMENTAL CONFIGURATION & ASSESSMENT

A. Data Collection & Testing Environments

System performance assessment required testing with 12 subjects (8 male, 4 females, 20–45 age spectrum). Each subject experienced recording under three different lighting situations:

- Daylight: Normal, adequately lit interior conditions
- Low-light: Night driving simulation with diminished surrounding illumination
- Backlight: Driving toward luminous background simula- tion (e.g., dawn/dusk situations)

Each participant executed planned activities including standard driving positions with regular blinking, simulated drowsiness (prolonged eye closure), deliberate yawning, speech, and con- siderable head movements. Approximately 3 hours of video material accumulated collectively. Manual frame-by-frame video labeling produced reference datasets for evaluation

B. System Interface and Results

Figures 2, 3, 4, and 5 exhibit Driver Safety Monitor system interfaces during different operational modes. The interface delivers live feedback comprising:

- Current EAR and MAR readings with color-coded dis- plays
- System condition (Active/OPERATIONAL/Inactive)
- Live statistics with average ratios
- Session overviews with incident counts
- Recent alerts with time records and severity levels
- Modifiable detection settings

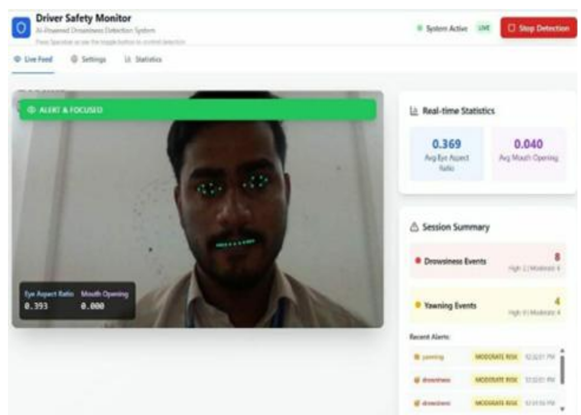

Fig. 2: Interface showing ALERT & FOCUSED con- dition with standard EAR (9.393) and MAR (0.000) readings. Session overview displays historical drowsiness

Fig. 3: YAWNNING DETECTED alert with MAR read- ing 0.120 surpassing threshold. Live statistics show nor- mal EAR (0.277) and raised MAR.

C. Assessment Criteria

System evaluation used standard classification measures:

- Precision: Correct positive identification proportion among all positive identifications (TP / (TP + FP))
- Recall (Sensitivity): Genuine positive instance propor- tion accurately detected (TP / (TP + FN))
- F1-score: Harmonic mean of precision and recall, sup- plying balanced measure (2 * (Precision * Recall) / (Precision + Recall))
- Latency: Average time between reference incident initi- ation (e.g., eye closure) and buzzer activation

D. Outcomes

Combined system performance across all subjects and con- ditions appears in Table I.

The system attained elevated F1-scores of 0.94 for drowsi- ness identification and 0.90 for yawning detection, confirming performance. Average complete system delay measured 75 ms with 18 ms standard deviation. Operation preserved strength in daylight and intermediate low-light environments but en-

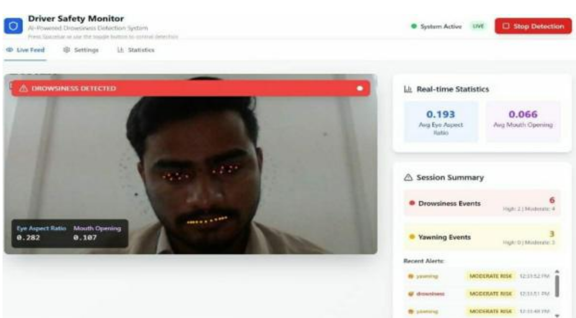

Fig. 4: DROWSINESS DETECTED alert with EAR reading 0.282. System presents live statistics and session overview with recent alert history.

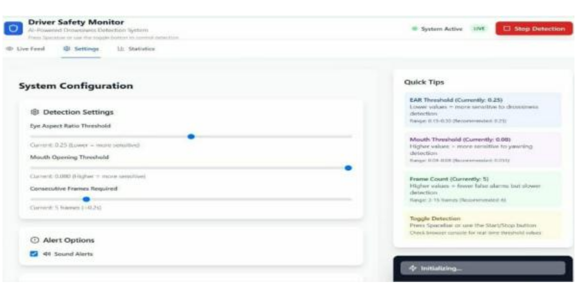

Fig. 5: Adjustment panel showing modifiable EAR (0.25), MAR (0.080), and frame number (5) limits with sensitiv- ity instructions.

Table I: Detection Performance (Combined)

Incident Category	Precision	Recall	F1-score
Drowsiness (EAR)	0.93	0.95	0.94
Yawning (MAR)	0.90	0.89	0.90
Overall Detection	0.92	0.93	0.925

countered minor decline in extreme backlight scenarios where facial characteristics were partly invisible.

VII. ANALYSIS

A. Reliability and Error Patterns

MediaPipe's Face Mesh algorithm supplied consistent landmark positioning even with intermediate head rotations (roughly 30 degrees horizontal and vertical). Temporal filtering methodology demonstrated vital for minimizing false alarms from quick blinking or momentary wide mouth openings during conversation.

Principal false alarm origins incorporated extreme head angles or short-term blockages (e.g., hand covering mouth or eyes). False negatives primarily occurred when subjects used sunglasses, entirely obstructing eye landmarks. This con-

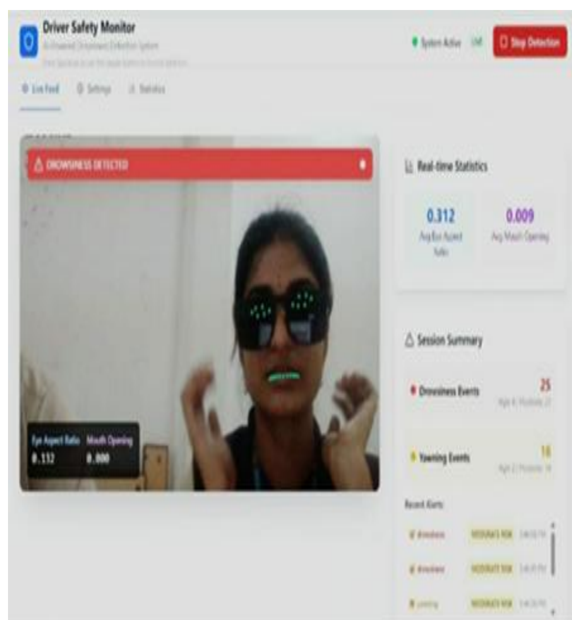

Fig. 6: DROWSINESS DETECTED alert displaying EAR0.167 and mouth aperture 0.113. System shows live statistics with mean EAR (0.239) and MAR (0.007), together with session overview indicating 1 severe and 6 intermediate drowsiness incidents.

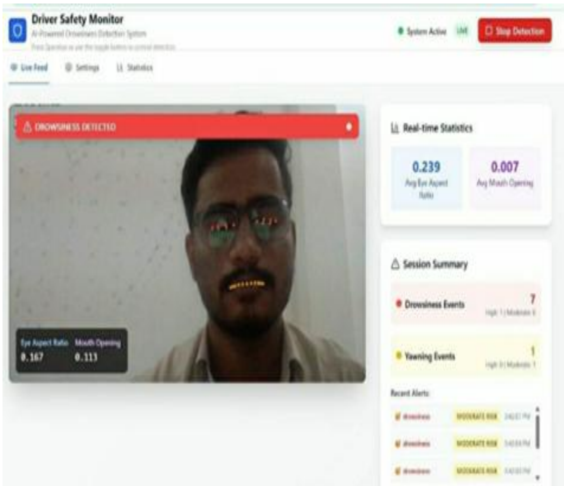

Fig. 7: DROWSINESS DETECTED condition with ex- tremely low EAR reading 0.132 denoting significant eye closure. System exhibits complete session statistics with 25 drowsiness and 16 yawning incidents identified. stitutes an acknowledged limitation of vision-based systems missing infrared imaging facilities.

B. Contrast with Deep Learning Methods

While comprehensive deep learning models can achieve leading-edge accuracy, they commonly require substantial computational assets (GPUs) and extensive, labeled train- ing datasets [10]. Our EAR/MAR-based approach provides a compelling option through computational economy, high comprehensibility, and removed training necessities, rendering it exceptionally suitable for low-capacity embedded processor implementation.

C. Restrictions

Primary system constraints incorporate:

- Blockage Performance: System malfunction happens when operator eyes or mouths are hidden (e.g., sunglasses or face coverings)
- Illumination Vulnerability: Performance decline under extreme low-light or powerful backlight conditions where landmark identification consistency reduces
- Contextual Comprehension Deficiency: System re- liance on geometric signals prevents discrimination between fatigue-caused yawning and boredom-related yawning, although both justify alerts

VIII. CONCLUSIONS AND PROSPECTIVE RESEARCH

This paper demonstrated a real-time, cost-efficient, and capable driver drowsiness and yawning identification system. By combining MediaPipe's high-capability facial landmark detection with validated EAR and MAR metrics and Arduino-controlled alert systems, we created a functional solution fitting for actual driving condition deployment. The system ex- hibited high identification accuracy and minimal delay across assorted realistic situations.

Future investigation directions incorporate:

- Multi-modal Combination: System strength improve- ment via extra data stream inclusion, such as head- position assessment (nodding detection) or vehicle signals from CAN bus (steering wheel angle)
- Infrared Vision: Infrared camera and LED incorporation would permit reliable functioning in total darkness and resolve sunglasses obstruction (IR permeable)
- Streamlined ML Algorithms: Efficient machine learn- ing models (compact neural networks or support vector machines) educated on EAR/MAR time-series informa- tion could learn intricate fatigue configurations, possibly improving decision logic beyond elementary thresholding
- Sophisticated Alert Systems: System outputs could combine with advanced driver-assistance systems (ADAS) for measures like seat or steering wheel vibration, or automatic emergency system initiation in crucial circumstances

IX. ACKNOWLEDGMENT

The investigators acknowledge the Department of Artificial Intelligence and Machine Learning at Sri Vasavi Engineering College, Tadepalligudem, for supplying essential laboratory equipment and assistance for this research. The investiga- tors also communicate sincere thankfulness to Mrs. Balaji Rohitha for her exceptional direction, advisory support, and persistent motivation throughout this research project.

REFERENCES

[1] World Health Organization," Global Status Report on Road Safety," WHO, 2023.

[2] S. Lal and A. Craig," Driver Fatigue: Electroencephalography and Psychological Assessment," Psychophysiology, vol. 38, no. 2, pp. 298- 306, 2001.

[3] K. Fujiwara et al.," Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG," in IEEE Transactions on Biomedical Engineering, vol. 66, no. 6, pp. 1769-1778, June 2019.

[4] S. Kim," Comprehensive Review on Driver Fatigue Detection Tech- niques," IEEE Access, vol. 10, pp. 12321-12338, 2022.

[5] M. Eriksson and N. Papanikotopoulos," Driver Fatigue Detection through Steering Wheel Movements," Proc. IEEE ITSC, 2001.

[6] M. Miyaji, M. Danno, and T. Maeda," Analysis of Driver Behavior Based on Lane Deviation," Proc. IEEE IVS, 2008.

[7] T. Soukupova´ and J. Cˇ ech," Real-Time Eye Blink Detection Using Facial Landmarks," 21st Computer Vision Winter Workshop, 2016.

[8] J. Jo, S. Lee, and K. Park," Vision-Based Eye Gaze Tracking and Yawning Detection," Sensors, vol. 11, pp. 8650-8669, 2011.

[9] M. Zhang," Driver Fatigue Detection System Based on Yawning Anal- ysis," Procedia Computer Science, vol. 110, pp. 266-272, 2017.

[10] B. K. Savas¸ and Y. Becerikli," Real Time Driver Fatigue Detection System Based on Multi-Task ConNN," in IEEE Access, vol. 8, pp. 12491-12498, 2020.

[11] P. S. Lamba et al.," Detection of Driver Drowsiness Using Adaptive Eye Characteristic Ratio for Enhanced Road Safety," in IEEE Access, vol. 13, pp. 84313-84328, 2025.

[12] Google MediaPipe," Cross-Platform ML Solutions," 2024. [Online]. Available: https://mediapipe.dev/