DRIVER SLEEP DROWSINESS DETECTION SYSTEM

Capstone-I project report submitted

by the student of

Hybrid UG program in Computer Science & Data Analytics

Shivam Kumar Singh

Roll No.: UA2503CDH566

Group No.: 98

INDIAN INSTITUTE OF TECHNOLOGY

PATNA BIHTA – 801106, INDIA

Date: 28/03/2026

Declaration

I hereby declare that this submission is my own work and, to the best of my knowledge, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Date: 28/03/2026

Student Name – Shivam Kumar Singh

Roll No – UA2503CDH566

Group No – 98

Signature – Shivam singh

Summary of the Project

This project presents the planning and implementation of a Driver Sleep Drowsiness Detection System using computer vision techniques. Driver drowsiness is one of the major causes of road accidents, and there is a strong need for a real-time, low-cost solution that can detect early signs of sleep and alert the driver before a critical situation occurs.

The proposed system uses a webcam to continuously monitor the driver's face and eye movements. It leverages OpenCV for video processing, Dlib for face and facial landmark detection, and the Eye Aspect Ratio (EAR) as the primary metric for determining whether the driver's eyes are open or closed.
In the final implementation stage, the system was further enhanced using Mouth Aspect Ratio (MAR/MER) analysis to detect yawning behavior as an additional fatigue indicator alongside EAR-based eye monitoring.

Project Github Repository: .link*

Team Details:

Name & Roll	Role	Responsibilities
Shivam Kumar Singh (UA2503CDH566)	System workflow design & Algorithm/Research Lead.	Responsible for designing system workflow, Researching and implementing EAR model and conducting mathematical validation with threshold calibration to ensure reliable and consistent performance.
Nikhil Kumar Singh (UA2503CDH703)	Backend (Main Code)	Responsible for the main backend implementation, including the primary detection logic and integration of computer vision components.
Priyanshu Kumar Singh (Group Member)	Backend	Responsible for backend-related tasks and supporting the integration of core logic and modules.
Ayush Kumar Singh (Group Member)	Frontend	Responsible for the frontend interface, ensuring that the user-facing components are clear and easy to use.
Rajnikant (Group Member)	JavaScript Styling	Responsible for JavaScript styling and visual enhancements using Three.js, contributing to the overall presentation and user experience.

Contents

1. Introduction

This project addresses a critical safety issue: driver drowsiness. Driver drowsiness is one of the major causes of road accidents. When drivers become sleepy or lose focus, their reaction time slows down, their awareness decreases, and the risk of collisions increases significantly.

This project focuses on detecting drowsiness in real-time using computer vision techniques. The system monitors eye movements and alerts the driver when signs of sleep are detected. By identifying drowsiness early and providing an immediate alert, the system aims to reduce the likelihood of accidents and improve road safety.
Our approach leverages widely available hardware—a standard webcam—and open-source libraries like OpenCV and Dlib. The core concept revolves around calculating the Eye Aspect Ratio (EAR) from eye landmarks. If the EAR remains below a threshold for a sustained period, the system identifies the driver as drowsy and triggers defensive alerts.

---

2. Problem Statement

Many accidents occur due to drivers falling asleep while driving. Traditional safety measures such as seat belts and airbags are reactive rather than proactive. There is a clear need for a real-time, low-cost system that can:

- Detect drowsiness accurately through non-invasive monitoring.
- Operate in real-time with low latency to ensure timely intervention.
- Provide immediate, unambiguous alerts to the driver.

The objective of this project is to implement an EAR-based monitoring system that can reliably distinguish between natural blinking and the onset of sleep.

3. System Overview

The proposed Driver Sleep Drowsiness Detection System functions as a real-time monitoring loop. It captures high-frequency video frames and processes them instantly to analyze eye states.

The system is built upon three technological pillars:

1. **OpenCV:** Handles high-performance video capture and frame pre-processing.
2. **Dlib:** Provides a high-fidelity 68-point facial landmark predictor (used specifically for the 12 landmarks defining the eyes).
3. **Eye Aspect Ratio (EAR):** A geometric formula used to objectively measure the degree of eye closure.
4. Mouth Aspect Ratio (MAR/MER): Detects yawning behavior by analyzing mouth openness using facial landmarks.

The system logic ensures that alerts are only triggered for prolonged closures, effectively ignoring normal blinking patterns.

4. Team Members & Work Distribution

The project is structured collaboratively, ensuring that each critical component—from algorithm research to frontend presentation—is handled by dedicated team members:

- **Shivam (Algorithm/Research lead):** Responsible for system workflow design, EAR model implementation, and mathematical validation with threshold calibration.

- **Nikhil Kumar Singh (Backend Lead):** Responsible for core detection logic, Dlib integration, and script orchestration.

- **Priyanshu (Backend Support):** Handles backend module integration and data flow.

- **Ayush Kumar Singh (Frontend):** Developing the user interface for monitoring and system status display.

- **Rajnikant (Visual Styling):** Implementing advanced visual components using Three.js for a modern user experience.

---

5. System Architecture & Working

The system is designed as a real-time computer vision pipeline optimized for execution on standard hardware. A webcam captures continuous video frames, which are processed sequentially.

Face detection is applied to identify the driver, followed by facial landmark extraction using Dlib's 68-point model. The eye region is isolated, and the Eye Aspect Ratio (EAR) is computed.

The EAR values are analyzed temporally across frames to distinguish between natural blinking and sustained eye closure. A decision engine applies threshold and frame-based conditions to classify the driver's state.

If drowsiness is detected, visual and audio alerts are triggered. This architecture ensures low latency, modular processing, and reliable real-time performance.
In the final implementation phase, the architecture was extended using Mouth Aspect Ratio (MAR/MER) analysis. Mouth landmarks are extracted and analyzed to measure mouth openness and detect prolonged yawning behavior. The system combines both EAR and MAR signals to improve overall fatigue detection reliability.

5.1 Detection Logic

The detection logic employs a temporal threshold. Because blinking is a natural and brief closure of the eyes, the system must distinguish it from drowsiness. We define a "Drowsy State" if:

Average(EAR_Left, EAR_Right) < 0.21 FOR > 15 Frames

When this condition is met:

- A visual overlay "DROWSY" is rendered on the screen.
- The computer's internal speaker triggers a 1000Hz alert tone.

6. Main Code & Supporting Scripts

The technical implementation consists of the following primary components:

6.1 Real-time Detection Scripts

The logic captured in the screenshots below outlines the frame processing loop and the landmark extraction process.

```
import dlib
import winsound
import time
from scipy.spatial import distance as dist
from imutils import face_utils

# ================= GLOBAL VARIABLES =================
current_ear = 0
drowsy_state = False

# ================= EAR FUNCTION =================
def calculate_ear(eye):
    A = dist.euclidean(eye[1], eye[5])
    B = dist.euclidean(eye[2], eye[4])
    C = dist.euclidean(eye[0], eye[3])
    return (A + B) / (2.0 * C)

# ================= MAR FUNCTION =================
def calculate_mar(mouth):
    A = dist.euclidean(mouth[13], mouth[19])
    B = dist.euclidean(mouth[14], mouth[18])
    C = dist.euclidean(mouth[15], mouth[17])
    D = dist.euclidean(mouth[12], mouth[16])
    return (A + B + C) / (2.0 * D)

# ================= THRESHOLDS =================
EAR_THRESHOLD_NORMAL = 0.23
EAR_THRESHOLD_WARNING = 0.20
EAR_THRESHOLD_CRITICAL = 0.15

MAR_THRESHOLD = 0.6
MAR_CONSEC_FRAMES = 15

FRAME_WARNING = 10
FRAME_CRITICAL = 20
```

Figure 6.1: Main initialization and EAR function definition. [link*]()

```
# ================= GENERATE FRAMES =================

def generate_frames():
    global current_ear, drowsy_state

    cap = cv2.VideoCapture(0)

    prev_time = 0
    count = 0

    while True:
        success, frame = cap.read()
        if not success:
            break
```

Figure 6.2: Video capture loop implementation. [link*]()

7. My Work:

My contribution to this project focused on four key areas: system workflow design, research and implementation of the Eye Aspect Ratio (EAR) model, implementation of Mouth Aspect Ratio (MAR) based yawn detection, and mathematical validation with threshold calibration.

I designed the system workflow by defining how data flows from input (video capture) to processing (face and landmark detection), analysis (EAR computation), and final output (alert generation). This provided a clear structure for the overall system.

Secondly, I conducted detailed research on the Eye Aspect Ratio (EAR), which serves as the core feature for detecting eye closure. I studied how EAR behaves under different eye states and verified that it remains relatively stable when the eyes are open but drops significantly during closure. I also analyzed its ability to distinguish between normal blinking and prolonged eye closure, making it suitable for drowsiness detection.
In the final implementation phase, I additionally worked on Mouth Aspect Ratio (MAR) based yawn detection. I studied mouth landmark behavior during normal speech and yawning conditions, and analyzed how mouth openness changes during fatigue-related yawning events. This was integrated as an additional fatigue indicator alongside EAR-based monitoring.

Finally, I performed mathematical validation of both EAR and MAR models and contributed to threshold selection. By analyzing EAR and MAR values across different conditions, I identified effective thresholds and temporal conditions that distinguish natural facial movements from fatigue-related behavior, thereby reducing false positives.

Overall, my work focused on designing the system's operational logic, validating its core detection mechanisms, and ensuring accuracy, robustness, and real-time performance.

(a) • System workflow design

(Input → Processing → Analysis → Decision → Output**)**

1. Input Layer

- Continuous video stream captured via webcam
- Frames processed sequentially in real time

2. Face Detection

- Detects presence of driver's face in each frame
- Ensures processing only when a face is visible

3. Facial Landmark Extraction

- Uses 68-point facial landmark model
- Extracts key eye coordinates

4. Temporal Analysis

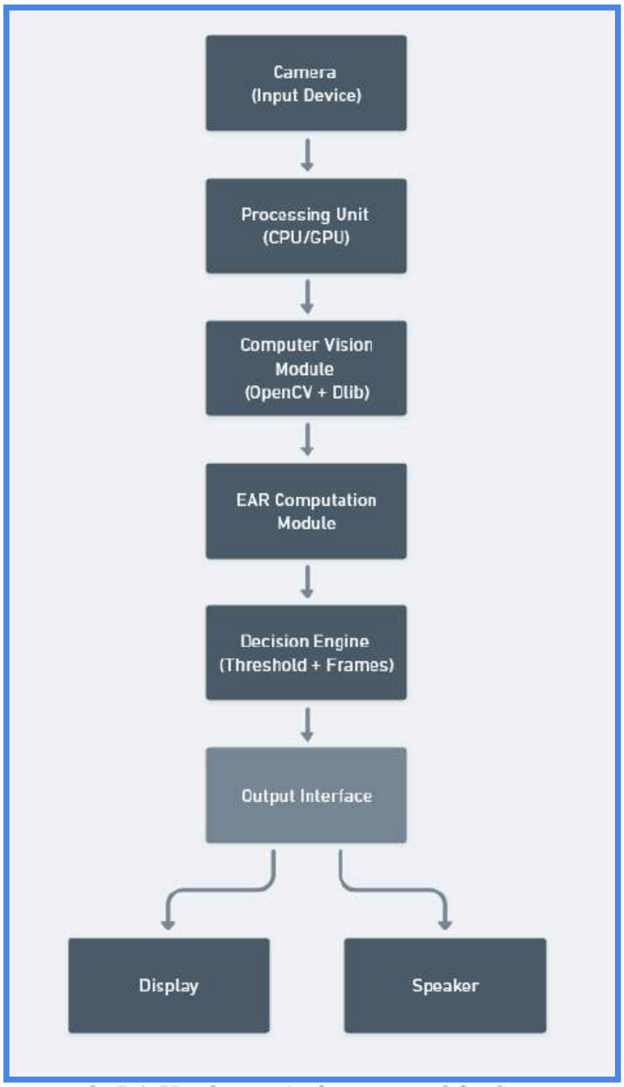
- Tracks EAR values across consecutive frames

5. Decision Engine Applies

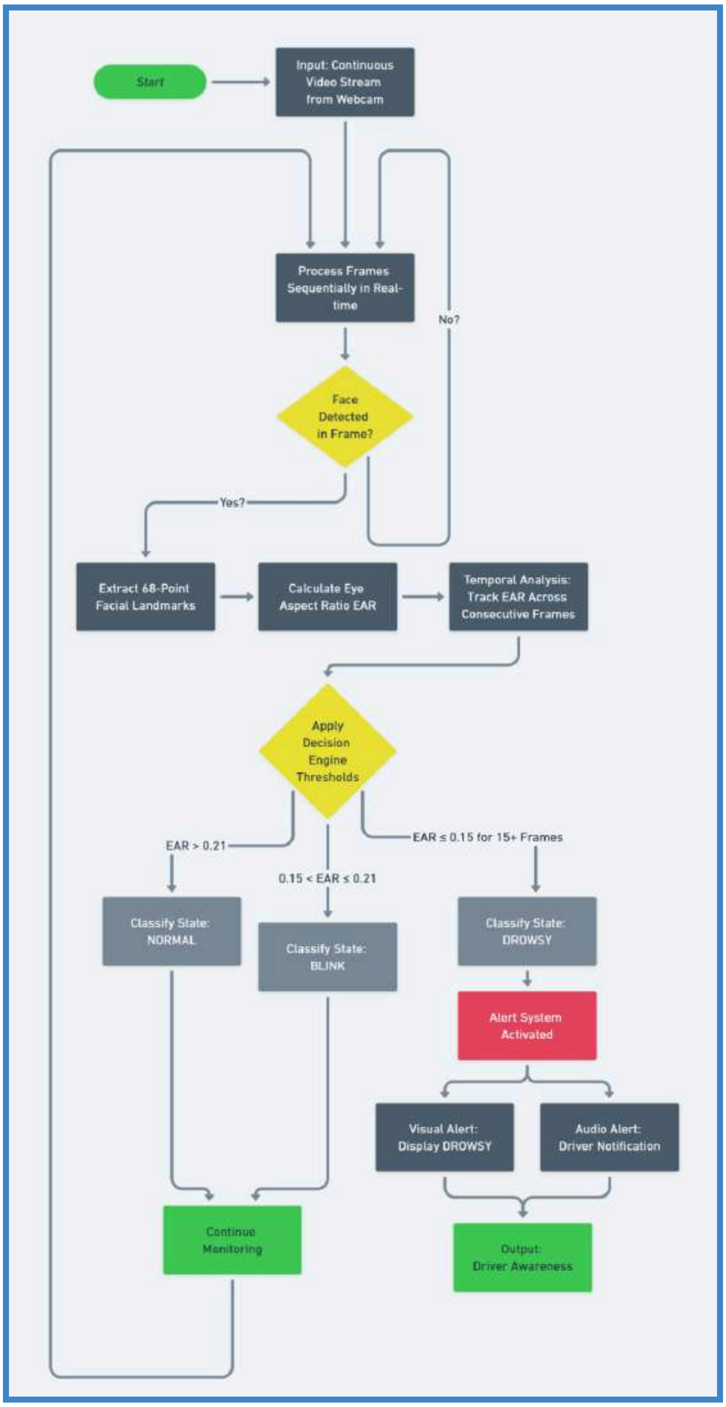
- EAR threshold (0.21)
- Frame condition (15 frames)
- Classifies state: Normal / Blink / Drowsy

6. Alert System

- Visual alert (“DROWSY”)
- Audio alert for driver awareness



- *fig7.1:Hardware Architecture of the System*



- *fig7.2:System Workflow Design*

(b) • EAR research & Implementation

The Eye Aspect Ratio (EAR) is a geometric measure used to detect eye openness based on facial landmarks.

1. EAR Formula:

EAR= (Vertical distance1+Vertical distance2) / (2*(Horizontal distance)

$$\boxed{EAR = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2||P_1 - P_4||}}$$

- **Working Principle**

Key Idea:

- Points p1-p6 define the eye
  P1-P4=vertical distance, P2-P6 , P3-P5=horizontal distance
- Vertical distances shrink when eyes close
- Horizontal distance remains stable

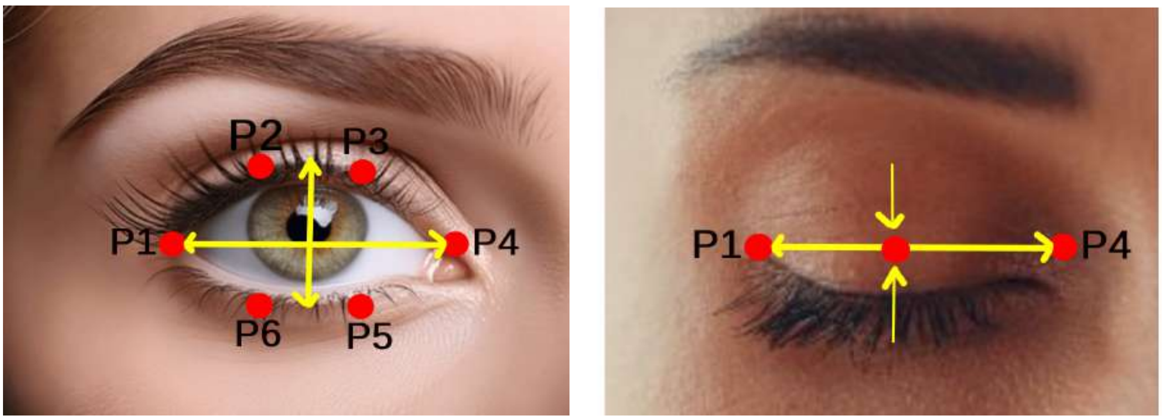


figure7.3: EAR landmarks diagram

figure7.4: EAR experiment observation table.

(c) • Mathematical validation & Threshold calibration

Objective

To ensure accurate classification between normal blinking and drowsiness using EAR.

Threshold Selection

- Open eye EAR ≈ 0.25–0.30
- Closed eye EAR ≈ < 0.20

Selected threshold: **0.21**

Temporal Logic

- Blink duration → 3–5 frames
- Drowsiness → sustained duration

Final condition:

EAR < 0.21 for more than 15 consecutive frames → Drowsy

Validation Approach

- Observed EAR values across frames
- Compared short vs long duration drops
- Verified consistency across conditions

Scenario	EAR Behavior	System Response
Eyes Open	Stable (~0.25–0.30)	No Alert
Blink	Short drop (<0.20)	Ignored
Drowsy	Sustained drop (<0.21)	Alert Triggered

Optimization Insights

- Combines spatial (EAR) and temporal (frames) analysis
- Reduces false positives
- Improves detection reliability.

Key Insight

The combination of threshold-based detection and temporal filtering ensures accurate classification, reduced false positives, and reliable real-time performance.

(c) • MAR Research & Implementation

The Mouth Aspect Ratio (MAR/MER) is a geometric measure used to detect mouth openness and yawning behavior using facial landmarks.

MAR=Horizontal Mouth Distance/Vertical Mouth Distance

🔹 Working Principle

Key Idea:

- Mouth landmarks define lip geometry
- Vertical mouth distance increases during yawning
- Horizontal distance remains relatively stable
- Sustained mouth opening indicates fatigue-related yawning

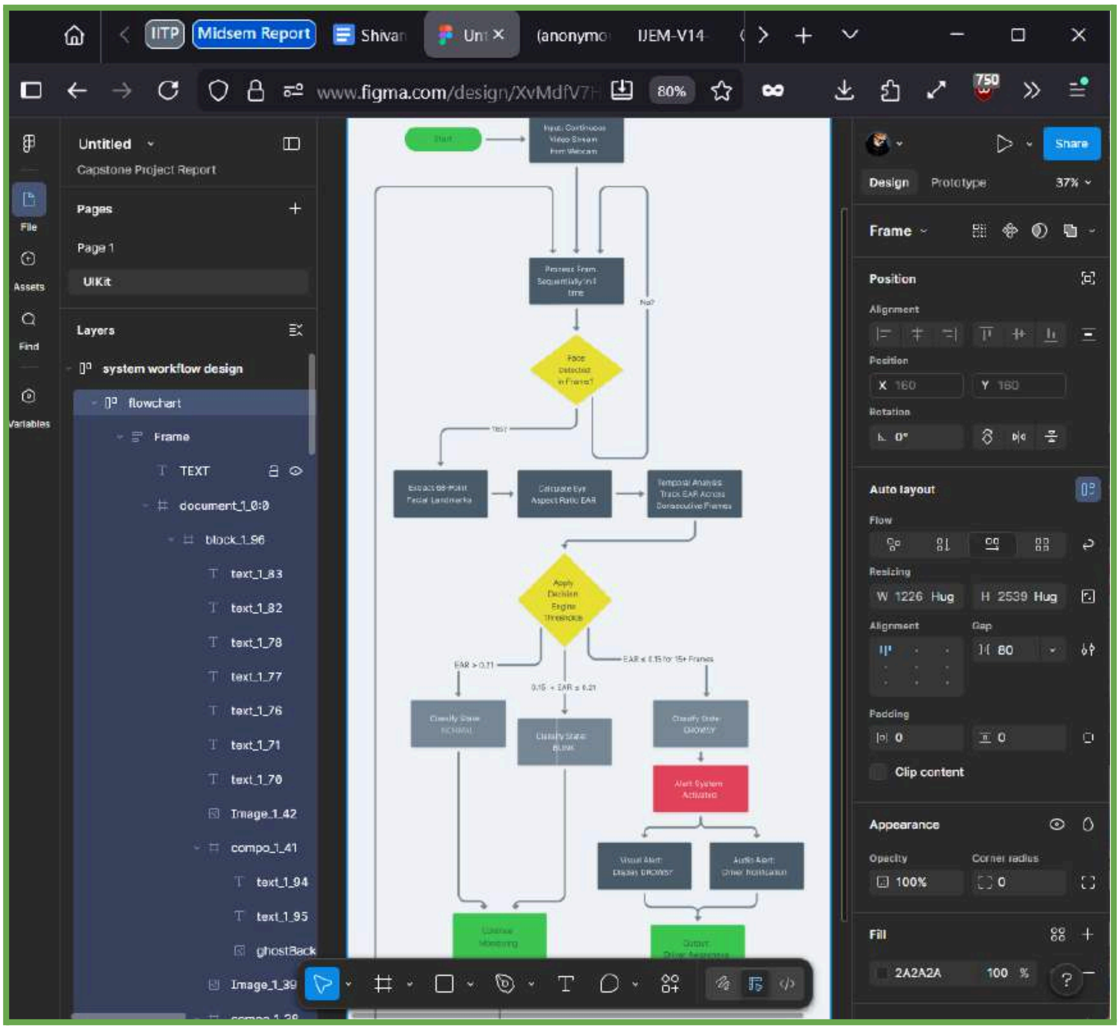
8. Proof of Work

My contributions are supported through design artifacts, observational analysis, and validation results.

8.1 System Workflow Design

The system workflow was designed using Figma, where the complete pipeline from input to alert generation was structured and refined.

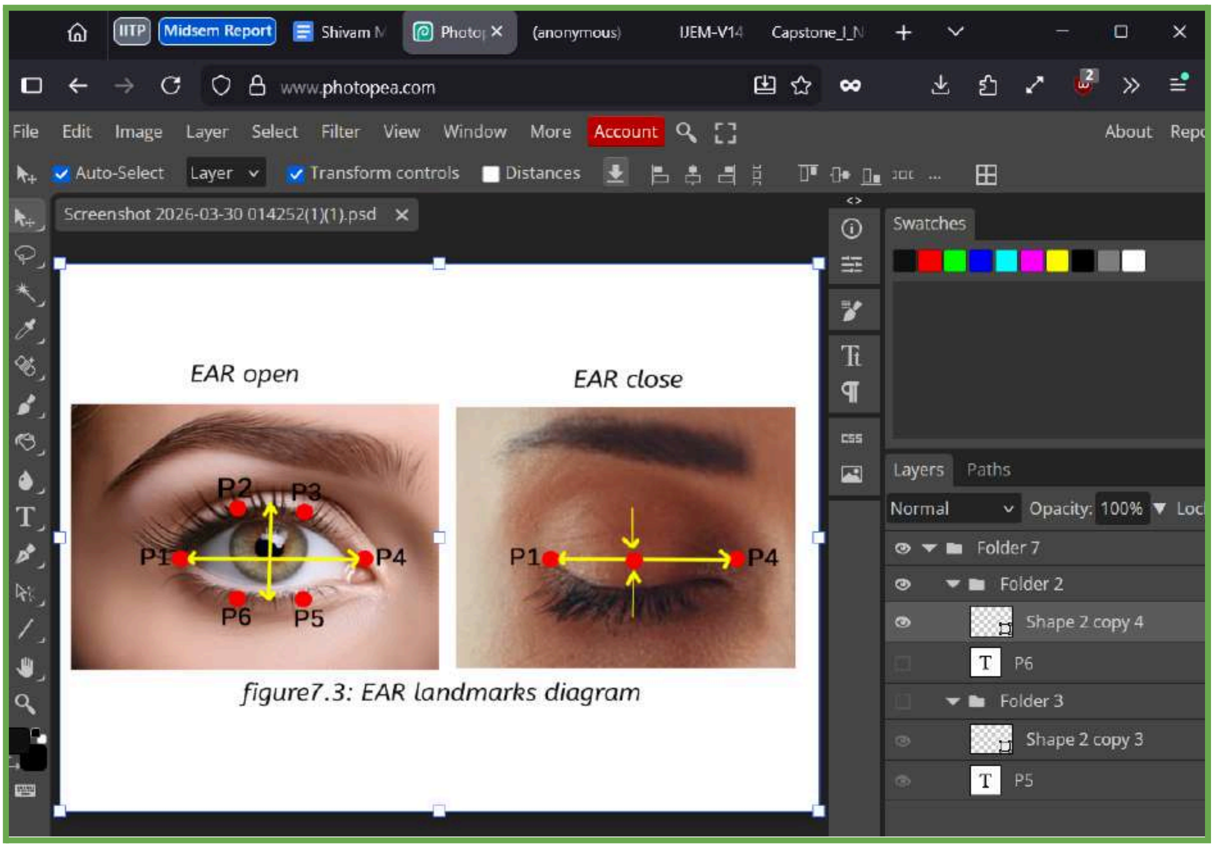
Evidence: Workflow diagram and Figma workspace screenshot (Figure 8.1)



8.2 EAR Research & Analysis

The EAR model was analyzed using landmark-based diagrams to understand its geometric behavior. It was observed that EAR remains stable for open eyes and decreases significantly during eye closure.

Evidence: EAR landmark diagram and editing proof (Figure 8.2)



8.3 Mathematical Validation & Threshold Calibration

EAR values were analyzed across frames to distinguish between blinking and drowsiness. A threshold of 0.21 with a temporal condition of 15 frames was validated as effective.

Evidence: EAR observation table and frame-wise analysis graph (Figure 8.3)

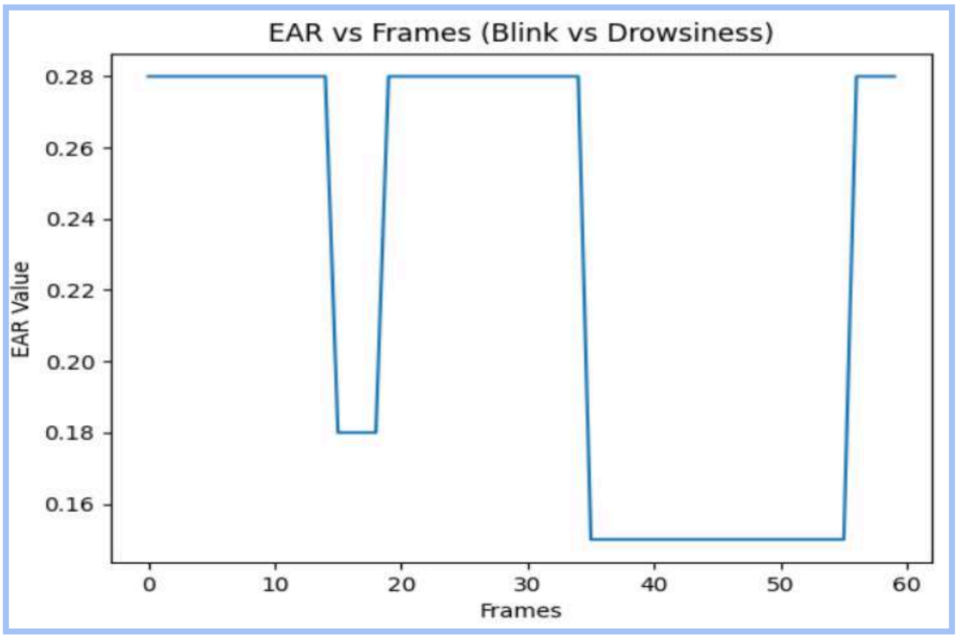


Figure 8.3: EAR variation across frames

Validation Outcome

- Accurate distinction between blink and drowsiness
- Reduced false positives using temporal filtering
- Stable real-time performance

10. Future Improvments

10. Future Work

While the final system is functional and capable of detecting both eye closure and yawning behavior, several enhancements can further improve robustness and scalability:

• Low-Light Optimization: Improving detection reliability under poor lighting conditions.

• Head Pose Estimation: Integrating head orientation tracking for enhanced attention analysis.

11. Conclusion

The Driver Sleep Drowsiness Detection System successfully demonstrates that computer vision can provide a robust and non-invasive solution for road safety. By utilizing facial landmark detection along with EAR and MAR analysis, the system effectively detects both prolonged eye closure and yawning behavior as indicators of driver fatigue.

The integration of temporal filtering and threshold-based validation ensures reliable real-time performance with reduced false positives. This project establishes a strong foundation for future intelligent driver monitoring and automotive safety systems.

12. References

1. . Real-Time Eye Blink Detection using Facial Landmarks (Soukupová & Čech,).pdf
2. Dlib Official C++/Python Library Documentation (Davis King).
3. OpenCV-Python Documentation (Bradski et al.).