Mediapipe Driver Monitoring Assignment

Amir Sasanfar - S323760

15. April 2025

1 Introduction

This assignment aims to develop a basic Driver Monitoring System in python, capable of detecting driver drowsiness and distraction by analyzing their head and eye gaze directions using an AI model. Based on this information, the program prints an alarm message when critical thresholds are exceeded, indicating that the driver is drowsy or distracted.

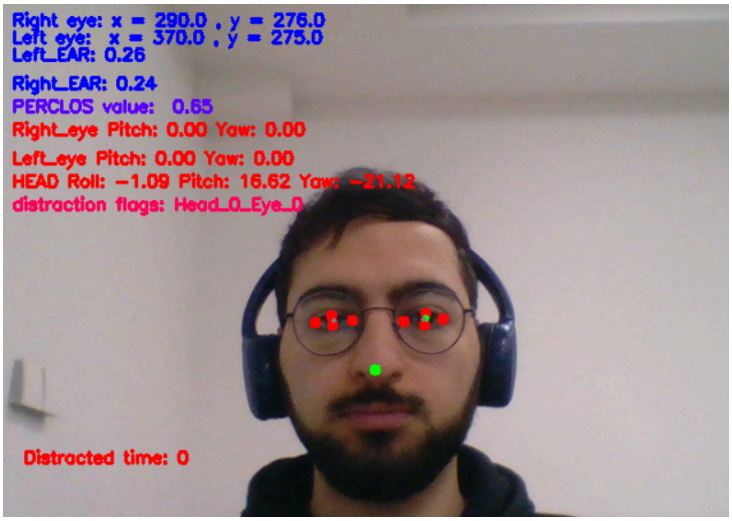

Figure 1: The interface of the final version of the Driver Monitoring system

1.1 Tools & Libraries

The system utilizes several tools and libraries: `OpenCV` (`cv2`) for video processing and drawing annotations, `MediaPipe` for facial landmark detection and gaze estimation, `NumPy` for numerical computations, and `time` and `statistics` for timing and signal processing tasks.

2 Development Details

This section outlines the key components implemented in the system for detecting drowsiness and distraction using facial landmarks and gaze direction analysis.

2.1 Drowsiness Detection

Drowsiness is identified using the Eye Aspect Ratio (EAR) and Percentage of Eye Closure(PERCLOS), calculated from specific eye landmarks. If the average EAR stays consistently outside a normal range, the driver is identified as drowsy.

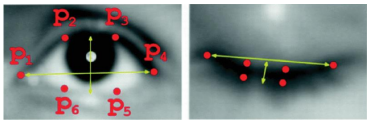

Figure 2: EAR for a normal open eye is between 0.3 and 0.4

2.1.1 EAR

The implementation extracts facial landmarks for both eyes and calculates the EAR for each eye separately, then averages them. A baseline EAR value of 0.3 is used as reference for a normal open eye, with drowsiness signaled when the EAR reamains above 80% of this baseline value for a sustained period. The EAR value is calculated using the function `ear_cal(coord_list: list)`

$$\text{EAR} = \frac{|Y_2 - Y_6| + |Y_3 - Y_5|}{2.|X_1 - X_4|} \tag{2.1}$$

Figure 3: Drowsiness Alarm appearing when the average EAR $> 80\%$ for more than 10 seconds

2.1.2 PERCLOS

PERCLOS (Percentage of eye Closure) is a standard measure used in driver monitoring systems. It represents the proportion of time the eyes are closed beyond a certain threshold within a given time window. The implementation maintains a history of EAR values over a 20-second sliding window and calculates the percentage of time the eyes are considered closed (when EAR falls below 0.25) using the function `perclos_cal(ear_history, threshold)`. And a `possible drowsiness alarm` is trigered when the threshold of 0.8 is crossed. As seen in Figure 4 the PERCLOS value is shown on the screen at all times. It is worth mentioning that my implementation of PERCLOS calculation deviates from that of the assignment description but is still a relevant approximation. Here a more statistical approach is considered.

$$\text{PERCLOS} = P(M, \theta) = \frac{|\{(e_i, t_i) \in M : e_i < \theta\}|}{|\text{M}|} \tag{2.2}$$

where:

- $M = \{(e_1, t_1), (e_2, t_2), ..., (e_n, t_n)\}$ is the set of EAR measurements and their timestamps
- θ is the threshold below which eyes are considered closed
- $|\{(e_i, t_i) \in H : e_i < \theta\}|$ counts measurements where EAR is below the threshold
- $|\text{M}|$ is the total number of measurements

Or put simply:

$\text{PERCLOS} = \frac{\text{count of measurements where eye closure}<\theta}{\text{Total number of measurements}}$ with the standard threshold of P80 for the ratio.

Figure 4: Alarm triggered when PERCLOS crosses the 80% threshold

2.2 Distraction Detection

Distraction is detected by analyzing the combination of head position and eye gaze direction, with the premise that a significant deviation from the forward-facing position indicates driver distraction.

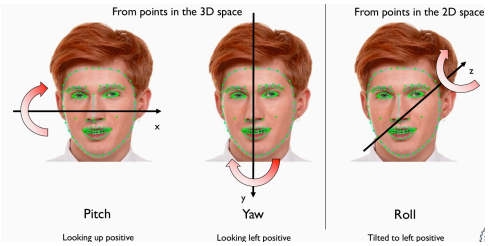

Figure 5: roll, pitch, yaw angles

2.2.1 Head Gaze Calculation

Head orientation is calculated using 3D facial landmarks and the camera matrix through OpenCV's solvePnP function, which provides rotation vectors. These are then converted to Euler angles (pitch, yaw, and roll). The head is considered distracted if any of these angles (roll, pitch, or yaw) exceeds 30 degrees from the neutral position.

2.2.2 Eye Gaze Calculation

Eye gaze direction is estimated by comparing the position of the iris center relative to the eye center. The system establishes a bounding ellipse for each eye and classifies gaze direction based on the iris position within this ellipse.

2.2.3 Combined Distraction Logic

Throughout the monitoring two `distraction flag variables` are used, one for head gaze and the other for eye gaze. As evident in Figure 6 both values along with a `distraction time` are continuously displayed on the screen and when both flags result in true(`equal to 1`) the distraction time is triggered. For the alarm to be less annoying and not be triggered for any negligible deviation a threshold of 200 ms was set for the distraction time condition. This value was chosen for testing a more standard value would be between 500 - 1000.

Figure 6: Distraction Alarm

3 References

1. Mediapipe
2. RidgeRun – PERCLOS for Driver Drowsiness Detection