Driver Monitoring System Using Emotion Detection

1. Introduction

This project implements a real-time driver monitoring system using computer vision and deep learning. The system analyzes facial expressions, eye and mouth movements, and head posture to detect emotional and physical states such as drowsiness, yawning, head nodding, and emotions like sadness, fear, and neutrality. Based on these inputs, the system can be integrated with an external control mechanism to adjust the speed of a cart , slowing down in unsafe conditions and maintaining speed when the driver appears attentive.

2. Project Goals

- Detect multiple faces simultaneously and analyze their states independently.

- Recognize driver tiredness indicators such as eye blinks, yawns, and head nodding.

- Detect and classify driver emotions in real-time using DeepFace and MediaPipe.

- Correct misclassifications (e.g., treat fear as neutral when indicators don't match).

- Provide a safety mechanism by interpreting emotional and physical states for vehicle speed control (hypothetically adjusting cart speed).

3. Technologies Used

- OpenCV : For real-time webcam video capture and frame processing.

- MediaPipe : For precise face landmark detection (eyes, mouth, nose).

- DeepFace : For emotion analysis using a deep learning model.

- NumPy : For numerical operations like distance calculations.

- Python : The overall implementation language.

- Collections (deque, defaultdict) : For maintaining per-face history and state.

4. Core Functionality

a. Face Landmark Detection (MediaPipe)

- Tracks up to 5 faces simultaneously.

- Extracts key facial landmarks for eyes, mouth, and nose.

b. Emotion Detection (DeepFace)

- Captures a cropped face image and predicts the dominant emotion.

- Applies smoothing to avoid flickering or unstable predictions.

- Includes logic to override fear or sad when facial context contradicts it.

c. Tiredness Detection

- Eye Aspect Ratio (EAR) detects blinking or eye closure.

- Mouth Aspect Ratio (MAR) detects yawning.

- A sustained low EAR triggers a drowsiness warning.

- A high MAR triggers a yawning warning.

d. Head Nod Detection

- Compares vertical displacement of the nose tip across frames.

- Large vertical shifts indicate nodding, a sign of sleepiness.

e. Emotion Correction Rules

- Emotion 'fear' downgraded to 'neutral' if facial features are calm.

- Emotion 'sad' reclassified if mouth corners are raised or EAR is high.

- Uses a rolling emotion history to ensure stability before updating UI.

f. Visual Output

- Displays:

    o Real-time bounding box around each face.

    o Emotion label above the face.

    o Warnings such as "⚠ Drowsiness", "⚠ Yawning", or "⚠ Head Nodding".

    o Facial landmarks overlaid for debugging and visualization.

5. Application Scenario

This system can be used to monitor a driver or operator in any context where alertness and emotional control are critical. A connected vehicle or cart can:

- Slow down if the driver is drowsy, yawning, or frequently nodding.

- Maintain normal speed when the driver appears alert and emotionally stable.

- Integrate into autonomous navigation systems, e-scooters, or industrial carts.

6. Limitations

- Emotion misclassification in low lighting or poor face angles.

- DeepFace latency may affect real-time performance on low-end machines.

- Assumes that the driver's face is always visible to the camera.

- No speed control logic implemented in this script — would need to be integrated separately.

- No facial identity tracking — faces are handled by index, not identity.

- It is challenging to accurately differentiate between subtle emotions such as sadness vs. neutrality or fear vs. calm, especially when facial expressions are mild or ambiguous.

8. Conclusion

This system is a powerful prototype for monitoring emotional and physical tiredness in drivers. By combining facial landmark tracking (MediaPipe) and emotion recognition (DeepFace), it offers real-time monitoring and the foundation for an intelligent response system, ensuring safety, reducing accidents, and supporting human-centric automation.