Detecting Driver Drowsiness Using Image Processing Techniques

Nikhil Mishra - 210668, Dhruv - 210338, Rishi Poonia - 210851

November 10, 2024

IIT Kanpur

Table of Contents

Introduction

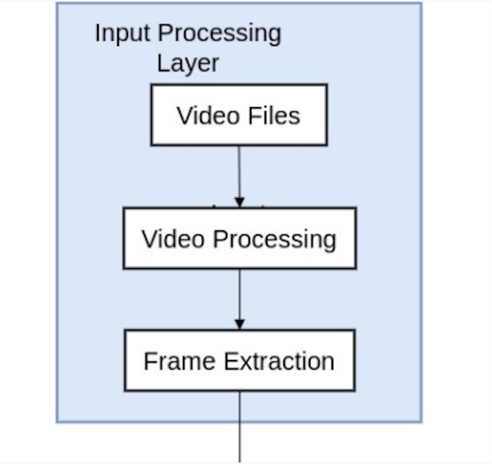
Introduction

- **Problem:** Driver drowsiness contributes to many road accidents.
- **Objective:** To develop a non-intrusive, image-processing-based system for detecting driver drowsiness.
- **Approach:** Analyzing facial landmarks and using deep learning for drowsiness detection.

Proposed Method

Overview of the Proposed Method

- **Steps:**
  - Video preprocessing
  - Feature extraction (EAR, MAR, PUC, MOE)
  - Classification using LSTM



Overview of the Proposed Method

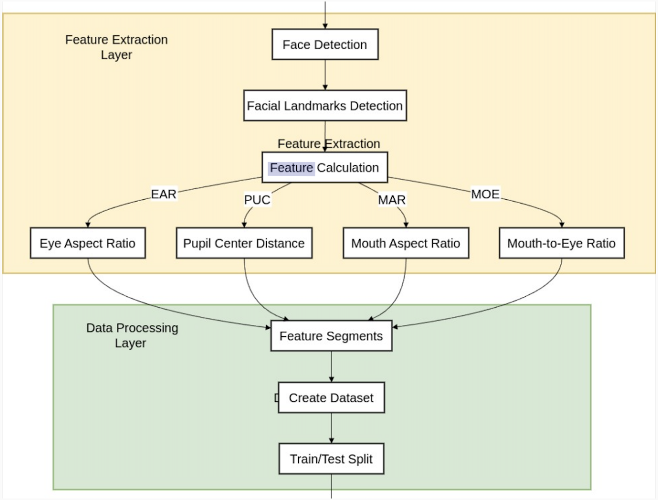


Figure 2: Overview of the Proposed Drowsiness Detection Method

Overview of the Proposed Method

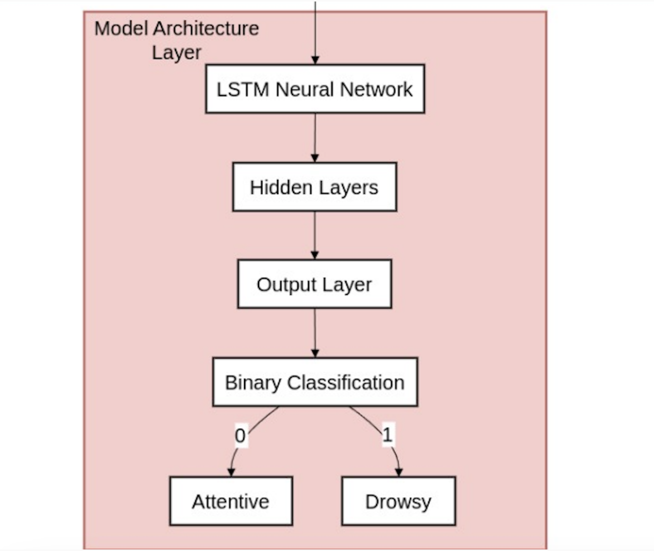


Figure 3: Overview of the Proposed Drowsiness Detection Method

Video Preprocessing

- **Objective:** Extract frames and detect facial landmarks to analyze drowsiness indicators.
- **Frame Extraction:** Videos are split into frames, allowing analysis of the driver's face over time.
- **Landmark Detection:** Using `dlib`, we detect 68 facial landmarks, focusing on eyes and mouth for feature calculations.
- **Library:** `dlib` provides pre-trained models for fast and accurate facial keypoint detection.
- **Output:** Processed frames with landmarks ready for feature extraction (EAR, MAR, PUC, MOE).

Feature Extraction

Feature Extraction: Ratios Used

- **Eye Aspect Ratio (EAR):** Measures eye openness to detect blinks or prolonged closures.
- **Mouth Aspect Ratio (MAR):** Quantifies mouth openness, indicating yawning.
- **Pupil-to-Eye Center Distance (PUC):** Tracks gaze stability.
- **Mouth-to-Eye Ratio (MOE):** Combines mouth and eye data for accuracy.

Figure 4: Facial landmarks used for feature extraction

Eye Aspect Ratio (EAR)

$$\text{EAR} = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2 \times ||p_1 - p_4||} \tag{1}$$

- **Purpose:** Detects eye blinks or sustained closures.

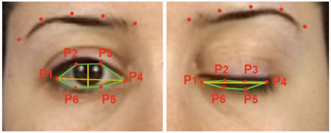


Figure 5: Eye landmarks: open vs. closed

Classification with LSTM

Classification with LSTM

- **Model Objective:** Use an LSTM network to analyze temporal patterns in facial features and classify drowsiness.
- **Input Features:** Segmented feature vectors (EAR, MAR, PUC, MOE) extracted from video frames.
- **Training Process:** LSTM learns sequential dependencies in features, improving accuracy in distinguishing between drowsy and alert states.
- **Output:** Classification of each segment as drowsy (1) or alert (0).

Experimental Results

Experimental Results

- **Dataset:** UTA RealLife Drowsiness Dataset
- **Metrics:** Precision, recall, F1-score, accuracy

	Precision	Recall	F1-Score	Support
0	0.82	0.82	0.82	114
1	0.78	0.78	0.78	94
Accuracy		0.80		208

Table 1: Classification Report of Best Model

Performance Metrics

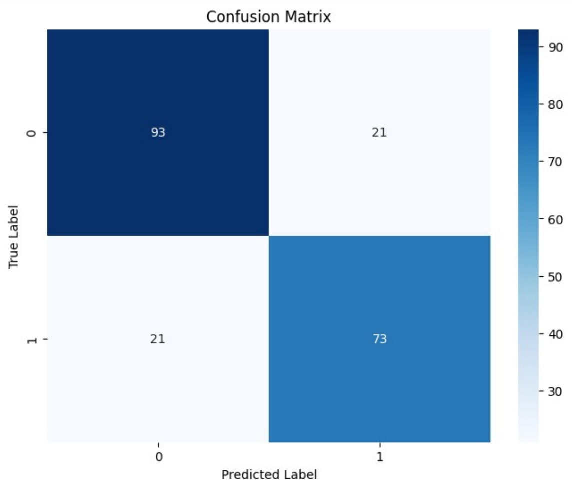


Figure 6: Confusion Matrix for Best Model

Performance Metrics

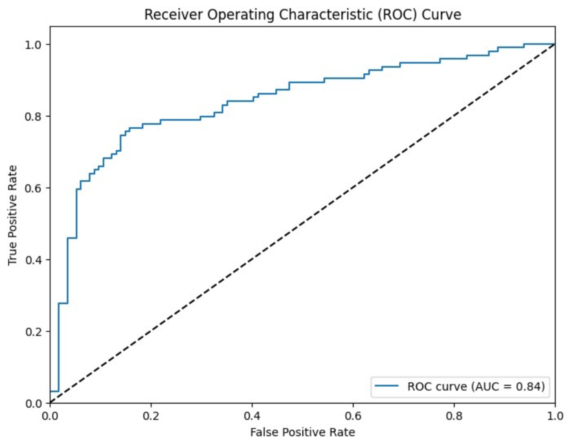


Figure 7: ROC Curve for Best Model

Experimental Analysis

Comparison of GRUNet and LightCNNLSTM

- **GRUNet:** Efficient with GRU layers, focuses on temporal dependencies.
- **LightCNNLSTM:** Combines CNN and LSTM for spatial and temporal feature extraction.

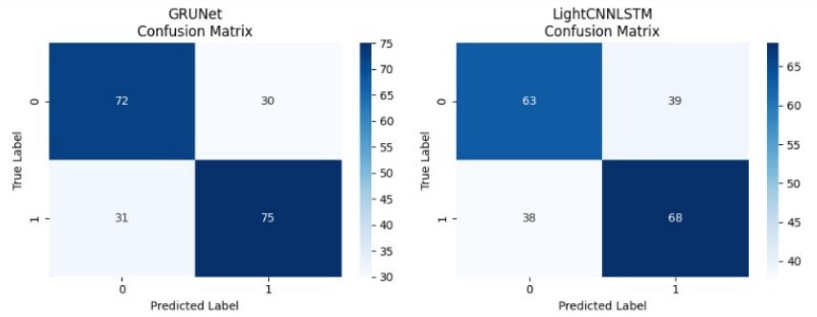


Figure 8: Confusion Matrix Comparison of GRUNet and LightCNNLSTM

Conclusion

Conclusion and Future Work

- **Conclusion:** Effective detection of driver drowsiness using facial landmarks and LSTM.
- **Future Work:** Real-time implementation and integration with vehicle systems.