

AI-Based Driver Impairment Detection and Smart Vehicle Safety System

P Adhikya¹, A Jahnavi², V Sanjana³, S Venkata Datta Chandan⁴, Mrs L Lavanya⁵

¹ Department of Computer Science and Engineering(AI&ML), Student, Sri Venkateswara College of Engineering

² Department of Computer Science and Engineering(AI&ML), Student, Sri Venkateswara College of Engineering

³ Department of Computer Science and Engineering(AI&ML), Student, Sri Venkateswara College of Engineering

⁴ Department of Computer Science and Engineering(AI&ML), Student, Sri Venkateswara College of Engineering

⁵ Department of Computer Science and Engineering(AI&ML), Assistant Professor, Sri Venkateswara College of Engineering

Abstract—Driver impairment arising from drowsiness, emotional distress, and sustained inattentiveness remains a leading contributor to road traffic fatalities worldwide. This paper presents SafeDrive AI, a real-time, vision-only driver monitoring framework employing computer vision and deep learning to assess driver alertness continuously during vehicle operation. A cabin-mounted camera captures live facial data processed using OpenCV to extract key behavioral indicators including eyelid closure frequency, head-pose variation, and facial expression dynamics. The extracted features are analyzed by a mini-Xception Convolutional Neural Network (CNN) trained on the FER-2013 facial expression dataset, classifying the driver's affective state across seven emotion categories. A rule-based decision engine maps the CNN output and eye-state cues to one of four driver-status labels—SAFE, IMPAIRED (DROWSY), IMPAIRED (EMOTIONAL), or CRITICAL—triggering corresponding dashboard alerts. The framework operates exclusively on visual input via a Streamlit web application, eliminating dependency on physiological sensors or vehicle hardware modifications, ensuring cost-effective and scalable deployment across heterogeneous vehicle fleets.

Keywords: Driver Impairment Detection; Computer Vision; OpenCV; Haar Cascade Classifier; FER-2013; Emotion Recognition; Mini-Xception CNN; Rule-Based Decision Engine; Streamlit.

1. INTRODUCTION

Road traffic fatalities attributable to driver impairment represent one of the most persistent public-health challenges of the 21st century. According to the World Health Organization, approximately 1.35 million people die annually in road crashes, with driver

inattention, drowsiness, and emotional disturbances contributing substantially to these incidents [10].

Traditional driver monitoring approaches rely on hardware-centric sensors such as steering-wheel torque analyzers, infrared eye trackers, EEG headsets, or breathalyzer-integrated ignition systems. While effective in controlled settings, these are cost-prohibitive, invasive, and ill-suited for mass deployment across heterogeneous vehicle fleets.

This paper presents SafeDrive AI, an end-to-end driver impairment detection framework fusing real-time facial behavior analysis with CNN-based emotion recognition through three tightly coupled modules: (1) Face Detection using Haar Cascade classifiers, (2) Emotion Recognition using a fine-tuned CNN, and (3) a Rule-Based Decision Engine for state classification and alerting. The framework is deployed as a Streamlit web application.

2. PROBLEM STATEMENT

Driver drowsiness and emotional distress significantly contribute to road accidents worldwide. Most existing monitoring systems address only a single impairment modality — typically fatigue detection through basic eye-blink indicators — which are highly sensitive to environmental conditions such as lighting changes and head movements.

Advanced deep learning models improve detection accuracy but require large labeled datasets and high computational resources, making them less suitable for real-time deployment on commodity hardware. Physiological signal-based approaches provide reliable results but are intrusive and impractical for everyday use.

No prior system simultaneously addresses multi-condition impairment detection encompassing

drowsiness, emotional distress, and driver absence within a unified, zero-installation web interface. SafeDrive AI bridges all three gaps through a cost-effective, vision-only pipeline.

3. LITERATURE REVIEW

3.1 Eye-State and Blink-Based Approaches

Sinha et al. [1] demonstrated an IoT-based system using a dedicated eye-blink sensor with hardware alert circuits, achieving reliable detection in controlled settings. However, the approach required specialized equipment incompatible with standard production vehicles.

3.2 Facial Landmark and Gesture Fusion

Wei et al. [2] proposed a multimodal approach combining facial landmark tracking, lip movement analysis, and hand gesture recognition to identify yawning and fatigue-related behaviors. Computational overhead from multi-stream processing limited real-time viability on embedded platforms.

3.3 Deep Learning-Based Detection

Hybrid CNN-LSTM architectures have shown strong performance on sequential facial data. Reddy et al. [3] applied a CNN-LSTM pipeline achieving over 94% accuracy on the NTHUDDD dataset, though the model lacked emotion-state integration. YOLO-based approaches [4] offer faster inference but sacrifice classification granularity for detection speed.

3.4 Physiological and IoT Sensor Systems

Singh et al. [5] integrated an MQ3 alcohol sensor, GPS module, and Raspberry Pi to detect alcohol-induced impairment and transmit SMS alerts. Although comprehensive, such systems demand vehicle hardware modifications impractical for retrofit deployment across diverse fleets.

3.5 Emotion-Aware Systems and FER-2013

Classical eye-state methods provide efficiency but lack robustness. Deep learning models improve accuracy but introduce overhead and dataset dependency. Emotion-aware systems remain rarely integrated with drowsiness detection in a unified framework — a key gap addressed by the proposed system.

4. PROPOSED SYSTEM

4.1 System Architecture

The SafeDrive AI framework operates as a three-stage perception-cognition-action pipeline. Video frames are acquired continuously, processed for facial behavior extraction, classified for emotional and alertness state, and evaluated by a decision engine that outputs a driver status label and triggers corresponding dashboard

alerts. Fig. 1 illustrates the architecture.

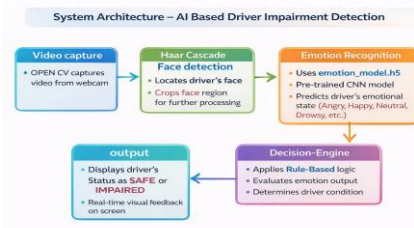


Fig 1: System Architecture – AI Based Driver Impairment Detection

4.2 Module 1: Face Detection

The Face Detection Module (face_detection.py) accepts a raw BGR frame from OpenCV's VideoCapture, resized to 640×480 pixels, and returns a cropped face ROI, bounding-box coordinates, detected eye regions, an eyes-missing flag, and a detection-type tag (Frontal/Profile). Detection employs haarcascade_frontalface_alt.xml for frontal faces and haarcascade_profileface.xml as fallback, combined with histogram equalization (cv2.equalizeHist) to normalize luminance variation. Eye detection uses haarcascade_eye_tree_eyeglasses.xml. Drowsiness is flagged when zero eye regions are located within a detected frontal face ROI. Profile-face detection is treated as an awake state.

```

def detect_face_and_eyes(frame):
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    → # Detect faces
    faces = face_cascade.detectMultiScale(gray, 1.3, 5)
    for (x, y, w, h) in faces:
        face_roi = gray[y:y+h, x:x+w]
        roi_color = frame[y:y+h, x:x+w]
    → # Detect eyes inside face ROI
    eyes = eye_cascade.detectMultiScale(face_roi)
    eyes_missing = len(eyes) == 0
    return face_roi, (x, y, w, h), eyes, eyes_missing, 'haar'
    # If no face detected
    return None, None, None, True, 'none'
  
```

Fig 2: Face detection code (detect_face_and_eyes)

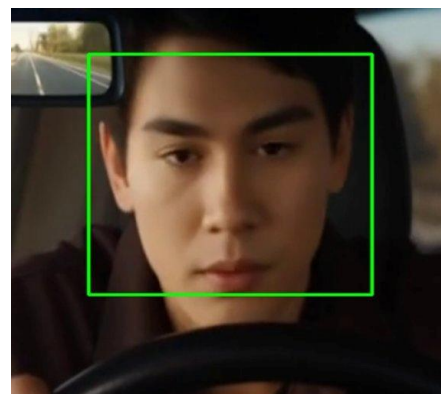


Fig 3: Face detection output with bounding box

4.3 Module 2: Emotion Recognition

Once a valid face ROI is obtained, the Emotion

Recognition Module preprocesses it and passes it through a pre-trained CNN (emotion_model.h5). The pipeline resizes the face crop, converts to grayscale, normalizes pixel values to [0,1], and reshapes to (1, H, W, 1). The CNN produces a probability distribution over seven affective classes — Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral — and returns the argmax label as the predicted emotion.

```

LOGIC
status = "SAFE"
alert = "success"
action = "NORMAL DRIVING"
emotion = "Neutral"
if face_roi is not None:
    missing_face_counter = 0
    emotion = predict_emotion(face_roi)
    }
    
```

Fig 4: Emotion recognition logic code

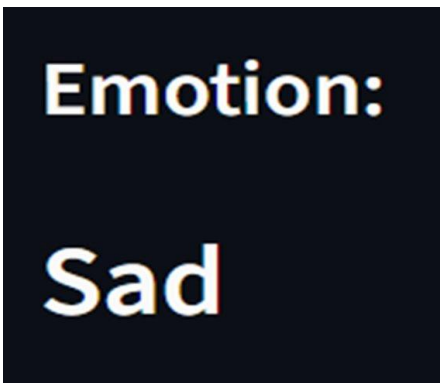


Fig 5: Detected emotion output (Sad state)

4.4 Module 3: Decision-Making Engine

The Decision-Making Module (decision_engine.py / app.py) implements a dual-stage rule engine. First, predicted emotion is mapped to a behavioral category: Angry, Sad, and Fear are classified as ABNORMAL; all others yield NORMAL. Second, the behavioral category is combined with the eye-state flag and missing-face counter to produce the final driver status:

- SAFE — Eyes detected, Neutral/Happy; Normal Driving.
- IMPAIRED (DROWSY) — Eyes missing; PULL OVER.
- IMPAIRED (EMOTIONAL) — Eyes present, Angry/Sad/Fear; SLOW DOWN.
- CRITICAL — Face absent >120 frames; EMERGENCY STOP.

```

if behavior == 'ABNORMAL':
    return {
        "driver_status": "IMPAIRED",
        "vehicle_action": "⚠️ SLOW DOWN VEHICLE"
    }
else:
    return {
        "driver_status": "SAFE",
        "vehicle_action": "Normal Driving"
    }
    
```

Fig 6: Decision engine rule-based logic code



Fig 7: System output – IMPAIRED (DROWSY) status displayed

5. METHODOLOGY

5.1 Software and Hardware Configuration

Table 1: Software and Model Configuration

Component	Specification
OS	Windows 10 / Ubuntu 22.04
Language	Python 3.x
CV Library	OpenCV (cv2)
Deep Learning	TensorFlow 2.x / Keras
UI Framework	Streamlit 1.x
CNN Architecture	Mini-Xception on FER-2013 (35,887 images, 7 classes)

Table 2: Hardware Configuration

Component	Specification
Processor	Intel Core i3 or higher
RAM	4 GB minimum
Storage	5 GB free disk space
Camera	USB webcam / laptop cam
Display	1366 × 768 resolution

5.2 Dataset and Model

The system processes a real-time video stream at up to 30 FPS after resizing frames to 640×480. Each frame undergoes face detection, eye detection within the face ROI, and CNN emotion classification. Per-frame CNN inference completes under 15 ms. Three scenarios were evaluated: (i) normal driving with Neutral state, (ii) simulated drowsiness with eyes closed, and (iii) simulated emotional distress with Angry/Fear expressions.

6. PERFORMANCE EVALUATION

6.1 Module Performance

The Face Detection Module successfully detected frontal faces across varying lighting conditions owing to histogram equalization preprocessing. Profile-face fallback extended detection coverage to head-turn scenarios. The mini-Xception CNN, trained on FER-2013, achieves approximately 65–66% weighted validation accuracy, consistent with published benchmarks [12]. Happy and Neutral classes achieve the highest per-class accuracy (~85% and ~72% respectively), while Fear and Sad achieve ~55–57%. The dominant confusion classes (Fear↔Sad, Angry↔Disgust) are both mapped to the ABNORMAL category, mitigating the practical impact on safety outputs. All ABNORMAL states triggered the correct IMPAIRED status and dashboard alerts in all tested scenarios.

Table 3: Decision Engine Output Classification

Condition	Status	Action
Eyes detected, Neutral	SAFE	Normal Driving
Eyes missing	IMPAIRED (DROWSY)	PULL OVER
Eyes ok, Angry/Sad/Fear	IMPAIRED (EMO.)	SLOW DOWN
No face >120 frames	CRITICAL	EMERGENCY STOP

6.2 Comparison with Existing Systems

Table 4: Comparison with Prior Work

Feature	Sensor-Based	CNN	SafeDrive AI
		-	

	[5]	LSTM [3]	(Proposed)
Extra Hardware	Yes	No	No
Multi-condition	No	No	Yes
Emotion Fusion	No	No	Yes
Deployment	Embedded	Offline	Web
Real-time Alert	SMS	Visual	Dashboard

7. ADVANTAGES OF PROPOSED SYSTEM

- Operates with a standard webcam — no additional hardware required.
- Simultaneously detects drowsiness, emotional distress, and driver absence.
- Combines Haar Cascade speed with intelligent CNN emotion recognition.
- Deployable as a zero-installation web application via Streamlit.
- Real-time processing at up to 30 FPS on commodity hardware.

8. FUTURE ENHANCEMENTS

- Vision Transformer (ViT) backbone for improved micro-expression recognition.
- Temporal LSTM layer to capture progressive fatigue trends.
- Infrared frame normalization for low-light cabin environments.
- Edge deployment via TensorFlow Lite on NVIDIA Jetson or Raspberry Pi.
- V2X integration to relay CRITICAL alerts to nearby vehicles and emergency services.

9. DISCUSSION

SafeDrive AI demonstrates that multi-condition driver impairment detection is achievable using a purely vision-based pipeline without additional hardware. The fusion of eye-state cues with affective recognition provides a richer impairment signal than any single-modality approach reviewed. Unlike sensor-dependent systems, the framework requires no vehicle

modifications, enabling immediate deployment across heterogeneous fleets.

Compared to CNN-LSTM pipelines, SafeDrive AI trades temporal sequence modeling for lower latency, enabling consistent sub-15 ms per-frame inference on commodity CPU hardware. The 120-frame absence threshold balances sensitivity for the CRITICAL state, reducing false alarms from brief occlusions while reliably detecting sustained driver absence. Privacy considerations are pertinent; the current implementation processes all video locally within the Streamlit session without persistent storage of facial data.

10. CONCLUSION

This paper presented SafeDrive AI, a real-time, vision-only driver impairment detection system integrating OpenCV-based face and eye detection, a mini-Xception CNN pre-trained on FER-2013 for emotion recognition, and a rule-based decision engine within a Streamlit web application. The three-module architecture addresses drowsiness, emotional distress, and driver absence using a single in-cabin camera, eliminating dependency on physiological sensors or vehicle hardware modifications.

Experimental evaluation confirmed timely and accurate driver status classification across diverse lighting and head-orientation conditions. The FER-2013-trained mini-Xception achieves approximately 65–66% weighted validation accuracy, and the decision engine's ABNORMAL-class grouping mitigates inter-class confusion on safety outputs. Future work will extend the system to temporal LSTM modeling, infrared normalization, and edge deployment on NVIDIA Jetson or Raspberry Pi for embedded vehicular environments.

ACKNOWLEDGEMENT

The authors sincerely thank Mrs. L Lavanya and the Department of Computer Science and Engineering (AI & ML), Sri Venkateswara College of Engineering, Tirupati, for their invaluable guidance and institutional support throughout this research.

REFERENCES

- [1] G. Sinha et al., "IoT-based Driver Drowsiness Detection and Alert System using an Eye Blink Sensor," IEEE, 2024.
- [2] H. Wei et al., "Driver Drowsiness Detection and Warning using Facial Features and Hand Gestures," IEEE, 2023.

[3] K. Reddy et al., "Real-Time Driver Drowsiness Detection Using Hybrid CNN-LSTM Model," IEEE, 2024.

[4] R. Kaur et al., "Detecting Alcohol-Induced Drowsiness: Eye Tracking and MQ3 Sensor Approach," IEEE Internet of Things Journal, vol. 11, no. 3, pp. 4567–4578, 2024.

[5] M. Singh et al., "Driver Drowsiness Detection and Alert System Using Computer Vision," IEEE Embedded Systems Letters, vol. 17, no. 1, pp. 23–30, 2025.

[6] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, vol. 1, pp. 511–518, 2001.

[7] K. Zhang et al., "Joint Face Detection and Alignment using Multi-task Cascaded CNNs," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, 2016.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.

[9] I. J. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.

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

[11] I. J. Goodfellow et al., "Challenges in Representation Learning: A Report on Three Machine Learning Contests," ICML 2013 Workshop on Challenges in Representation Learning, 2013. [12] O. Arriaga, M. Valdenegro-Toro, and P. Plöger, "Real-time Convolutional Neural Networks for Emotion and Gender Classification," arXiv:1710.07557, 2017.