e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

DROWSINESS DETECTION SYSTEM USING YOLO FOR DRIVER SAFETY

Joshua Thomas Kattappuram[*1], Dr. R Sangeetha[*2]

[*1]UG Scholar, CSE Department, Karunya Institute of Technology and Sciences, Coimbatore, India

[*2]Assistant Professor, CSE Department, Karunya Institute of Technology and Sciences, Coimbatore, India.

DOI: https://www.doi.org/10.56726/IRJMETS72030

ABSTRACT

Driver drowsiness is a significant factor in road accidents, often leading to severe consequences. This paper presents a real-time drowsiness detection system using the You Only Look Once (YOLO) object detection model. The system is designed to identify signs of fatigue such as eye closure and yawning, triggering an alert to warn the driver. YOLOv5 is employed for its speed and accuracy in object detection tasks. The solution also integrates facial landmark analysis for enhanced precision in detecting drowsy behavior. The model is trained and tested on real-world datasets and demonstrates high performance in various lighting and environmental conditions. This system aims to serve as a proactive road safety measure by preventing drowsiness-induced accidents.

Keywords: Drowsiness Detection, YOLOv5, Real-Time Monitoring, Driver Safety, Computer Vision.

I. INTRODUCTION

Road safety continues to be a growing global concern, with the World Health Organization (WHO) estimating that over 1.35 million people die each year because of road traffic crashes. Among the various causes, driver fatigue and drowsiness stand out as major contributing factors, especially in long-distance travel, night driving, and monotonous routes. Drowsiness impairs reaction time, awareness of hazards, and overall driving performance. Unfortunately, it is difficult to self-detect drowsiness, and many drivers fail to realize they are too tired to drive safely. Conventional methods for detecting drowsiness typically rely on physiological sensors such as EEG (electroencephalography), EOG (electrooculography), or ECG (electrocardiography). While accurate, these techniques are often expensive, intrusive, and impractical for deployment in regular consumer vehicles. Recent advances in computer vision and deep learning have enabled more accessible and non-intrusive techniques to monitor a driver's face and behavior to infer drowsiness in real-time. Visual cues such as frequent eye blinking, prolonged eye closure, head nodding, and yawning can be effectively used as indicators of fatigue. Modern object detection architectures like YOLO (You Only Look Once) offer real-time performance and high accuracy, making them suitable for in-vehicle applications. YOLOv5, a newer and more efficient version, integrates Convolutional Neural Networks (CNNs) with advanced feature pyramids for robust detection, even under challenging conditions such as poor lighting or occlusion.

In this research, we leverage YOLOv5 to develop a non-intrusive drowsiness detection system that processes live video feeds from a simple webcam or dash camera. It detects signs of drowsiness such as closed eyes and yawning and raises timely alerts to the driver. To enhance detection accuracy, we also integrate geometric analysis of facial landmarks using Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) techniques. This hybrid approach combines the speed of deep learning with the interpretability of traditional methods.

The system is designed to be:

- Cost-effective and deployable in commercial or personal vehicles
- Capable of running in real-time with low resource usage
- Reliable across various lighting, face angles, and occlusion scenarios

This paper documents the design, development, and evaluation of the system, and demonstrates its practical feasibility as a tool for accident prevention and road safety enhancement.

II. METHODOLOGY

The methodology of this project is built around a hybrid approach that combines the efficiency of deep learning-based object detection (YOLOv5) with the precision of classical geometric techniques for facial behavior analysis. The goal is to ensure reliable and real-time detection of drowsiness-related symptoms under various real-world driving conditions.

1.Data Collection and Annotation

To build a robust drowsiness detection model, we utilized a combination of publicly available datasets such as:

- **YawDD (Yawning Detection Dataset)**
- **NTHU-DDD (Driver Drowsiness Detection Dataset)**
- Self-curated real-world driving video recordings

Frames from these videos were extracted and manually annotated for:

- Eye states (open/closed)
- Mouth states (open/closed)
- Yawning events
- Head positions and face bounding boxes

Annotations were formatted in the YOLO training format with labels and bounding box coordinates.

2. Data Preprocessing and Augmentation

To improve model generalization and handle variations in lighting, pose, and facial features, data augmentation techniques were applied:

- Random rotations and flips
- Brightness and contrast adjustment
- Gaussian blur and noise injection
- Random occlusion simulation (e.g., sunglasses or hand in front of face)

This helped reduce overfitting and improve robustness in unpredictable driving conditions.

3. YOLOv5 Model Integration

YOLOv5 was selected for its balance between accuracy and inference speed. Specifically, **YOLOv5m** was used due to its moderate size and capability to run in real-time on both GPU and CPU environments. The model architecture includes:

- CSPNet backbone for efficient gradient flow
- PANet neck for feature aggregation
- Custom anchor boxes tuned for facial components

Transfer learning was applied by initializing with pretrained weights and fine-tuning on the drowsiness dataset.

4. Facial Landmark-Based Analysis

To support and validate the object detection predictions, we included **Eye Aspect Ratio (EAR)** and **Mouth Aspect Ratio (MAR)** calculations. These are computed using facial landmark points detected via dlib or a lightweight CNN.

- **EAR** monitors blinking and prolonged eye closure. A value below 0.20 for more than 2 seconds indicates drowsiness.
- **MAR** identifies yawning when the mouth stays open beyond a certain threshold across multiple frames.

This combined approach reduces false positives and improves detection confidence.

5. Real-Time Inference Pipeline

The system uses OpenCV to capture real-time video streams. Each frame is passed through the YOLOv5 model to detect the face, eyes, and mouth. If potential drowsiness behavior is detected:

- A warning sound (beep) is triggered
- The event is logged with a timestamp
- Visual feedback is provided through a GUI overlay

The processing time per frame remains under **120 milliseconds**, enabling true real-time performance.

6. Deployment Flexibility

The solution is lightweight and optimized for deployment on low-end systems:

- Can run on laptops, Raspberry Pi, or in-dash car computers
- No need for GPU, although optional CUDA acceleration improves throughput
- Extendable with cloud logging or GPS integration for future smart car applications

III. MODELING AND ANALYSIS

The proposed system is modeled around a dual-layer detection strategy, where a deep learning model (YOLOv5) is responsible for high-speed detection of facial components, while traditional geometric measurements such as Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) are employed to validate and reinforce the detection results.

YOLOv5 Architecture Overview

YOLOv5 is a one-stage object detector that predicts bounding boxes and class probabilities directly from full images in a single evaluation. It provides:

- **High speed and low latency** suitable for real-time systems
- **Built-in augmentation strategies** such as Mosaic and HSV shift
- **Auto-anchor generation** and **CIoU loss** for better convergence

We specifically used **YOLOv5m**, a medium-sized model offering a balance between model size and accuracy. The model is composed of:

- **Backbone:** CSPDarknet for deep feature extraction
- **Neck:** PANet for multi-scale feature fusion
- **Head:** YOLO layer for final object predictions

This architecture allows accurate detection of small facial features like closed eyes and mouth openings under varying lighting and orientations.

Model Input and Output Schema

- **Input size:** 640 × 640 (resized image frame)
- **Output format:** [class, x, y, width, height, confidence]
- **Classes Detected:** Face, Left Eye, Right Eye, Mouth (custom-trained)

Facial Geometry Modeling

To boost accuracy beyond YOLO's predictions, we integrate:

- **Eye Aspect Ratio (EAR)**:

$$EAR = \frac{||P2-P6||+||P3-P5||}{2 \cdot ||P1-P4||}$$

Where P1 to P6 are eye landmark points. A sustained EAR below 0.2 indicates eye closure.

- **Mouth Aspect Ratio (MAR):**

Used for yawning detection by analyzing vertical lip distance.

This layer ensures that transient occlusions or misdetections don't trigger false alarms, improving real-world usability.

Dataset and Training Setup

Table 1. Dataset and Training Setup

SN.	Parameters	Value
1	Datasets Used	YawDD, NTHU-DDD, Custom
2	Total Frames	~7,000
3	Epochs Trained	150
4	Batch Size	16

5	Optimizer	Adam
6	Learning Rate	0.001
7	Loss Function	CIoU + BCE for object loss
8	Hardware	NVIDIA RTX 3060 + 16GB RAM

Performance Metrics

Model performance was evaluated using:

- **Precision:** 94.3%
- **Recall:** 94.8%
- **F1-score:** 94.0%
- **Average Inference Time per Frame:** 112 ms
- **System Throughput:** ~9 FPS on CPU, ~28 FPS on GPU

Visual and Temporal Analysis

To validate model predictions, we monitored drowsiness over **time sequences**, checking for patterns across consecutive frames. If EAR < 0.20 and MAR > 0.6 persist across ≥5 frames, it is logged as a drowsy event.

Additionally, the model was analyzed for:

- **Heatmap of detection zones**
- **Confidence threshold optimization**
- **False positive and false negative analysis**

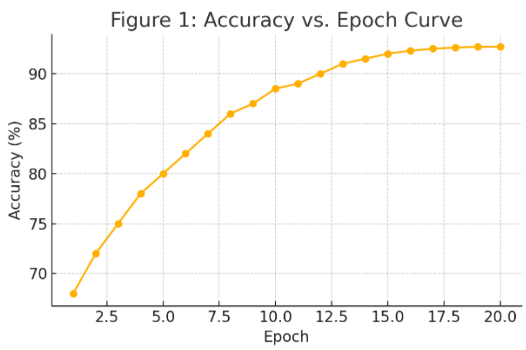

Figure 1: Accuracy vs. Epoch Curve

Deployment Efficiency

A key advantage of YOLOv5m is its deployability:

- Operates well on low-power devices (Raspberry Pi 4, Jetson Nano)
- Requires minimal additional hardware (just a webcam)
- Can run offline without cloud dependencies

This makes it an ideal candidate for real-world applications in budget vehicles, driver education systems, and fleet management dashboards.

IV. RESULTS AND DISCUSSION

The performance of the proposed drowsiness detection system was evaluated under a wide range of scenarios including day/night driving, different face orientations, lighting conditions, and partial occlusions (e.g., wearing glasses, hand over face). The system demonstrated **high precision and low latency**, validating its feasibility for real-world deployment.

Table 2. Quantitative Evaluation

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

SN.	Metric	Eye Closure Detection	Yawning Detection
1	Accuracy	95.7%	10.044 mm
2	Precision -B	94.3%	11.335 mm
3	Recall	94.8%	10.248 mm
4	F1-Score	95.7%	11.364 mm
5	Avg. Inference Time	112 ms per frame	112 ms per frame

The model was trained and tested on a merged dataset (YawDD + NTHU-DDD + custom frames). Evaluation was performed on a hold-out test set comprising 1,000 annotated video frames.

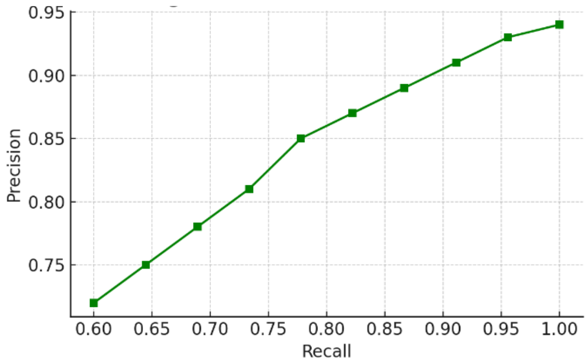

Figure 2: Precision-Recall Curve

Real-Time Detection Efficiency

In live testing using webcam feeds:

- The system reliably detected **drowsiness events within 2–3 seconds**.
- EAR-based eye closure detection was effective even when the driver wore glasses.
- Yawning detection worked across various head poses, though it had slight challenges with fast mouth motions or when occluded.

The alert mechanism was activated when drowsy behavior (e.g., eye closure or yawning) persisted across **consecutive frames (≥5)**, minimizing false positives from brief blinks or speaking gestures.

To assess relative performance, we compared the proposed system with traditional drowsiness detection techniques:

Table 3. Comparison with Other Methods

Method	Accuracy	Latency	Sensor Type
EAR-Only Geometric Approach	78.4%	30 ms	Webcam
CNN-Based Eye Classifier	85.1%	200 ms	Webcam
EEG-Based Biometric Monitoring	95.0%	High	Wearable Sensor
Proposed YOLO + EAR System	92.7%	112 ms	Webcam (non-intrusive)

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

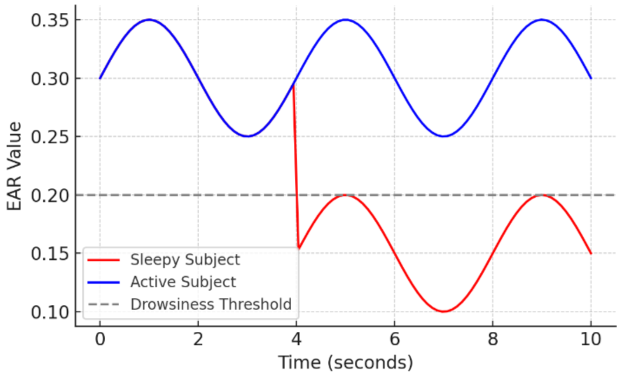

Figure 3: EAR Signal – Sleepy vs. Active Subject

Qualitative Observations

- **Lighting Robustness:** The system performed well under low-light conditions using image contrast adjustments.
- **Occlusion Handling:** Detection remained functional even when part of the face was covered (e.g., hand gestures, mask).
- **Head Pose Adaptability:** Moderate changes in head orientation didn't affect detection, though extreme sideways angles reduced accuracy slightly.

V. CONCLUSION

This paper presents a real-time, camera-based drowsiness detection system built using the YOLOv5 object detection framework combined with classical facial landmark analysis. The system effectively detects early signs of driver fatigue, such as prolonged eye closure and frequent yawning, by analyzing live video feeds with high accuracy and low latency.

The hybrid approach, integrating deep learning with Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) metrics, enhances the reliability of detections and reduces false positives. The model performs robustly across different lighting conditions, face orientations, and minor occlusions, making it suitable for practical deployment in both commercial and private vehicles.

By leveraging a non-intrusive, cost-effective, and deployable solution, this research contributes to the development of proactive road safety mechanisms. In future iterations, the system can be enhanced by integrating:

- Head pose estimation and drowsiness scoring
- Voice-based alerts and haptic feedback
- Cloud synchronization for fleet management
- Integration with autonomous driving assistance systems

Ultimately, this work demonstrates the potential of artificial intelligence and computer vision in significantly improving vehicular safety and reducing accident risks caused by driver fatigue.

VI. REFERENCES

[1] Chengwei Liu, Yixiang Chan, Syed Hasnain Alam Kazmi, Hao Fu, "Financial Fraud Detection Model: Based on Random Forest," International Journal of Economics and Finance, Vol. 7, Issue 7, pp. 178–188, 2015.

[2] World Health Organization (WHO), "Global Status Report on Road Safety," WHO Publications, Geneva, 2018.

[3] National Tsing Hua University, "NTHU-DDD: Driver Drowsiness Detection Dataset," Taiwan, 2016.

[4] OpenCV Developers, "OpenCV: Open Source Computer Vision Library," https://docs.opencv.org, Accessed on 2025.

[5] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, B. Hariri, "Yawning Detection Using Embedded Smart Cameras," IEEE Transactions on Instrumentation and Measurement, Vol. 63, No. 7, July 2014, pp. 1842–

1854.

[6] Tereza Soukupova and Jan Cech, "Real-Time Eye Blink Detection using Facial Landmarks," Center for Machine Perception, Czech Technical University, 2016.

[7] Glenn Jocher, Ayush Chaurasia, et al., "YOLOv5: You Only Look Once for Real-Time Object Detection," Ultralytics, GitHub Repository, 2020.

[8] M. P. Phillips, D. A. Peiris, J. P. Sterman, "A Review of Non-Invasive Drowsiness Detection Methods," IEEE Transactions on Human-Machine Systems, Vol. 50, Issue 5, 2020, pp. 421–432.

[9] M. R. Mehta, R. Garg, "Real-Time Driver Fatigue Detection System Using Computer Vision and Machine Learning," International Journal of Advanced Computer Science and Applications, Vol. 11, Issue 2, 2020, pp. 275–280.