Received: 9 September 2025 | Revised: 4 December 2025 | Accepted: 6 February 2026 | Published online: 5 March 2026

Journal of Computational and Cognitive Engineering
2026, Vol. 00(00) 1–11
DOI: 10.47852/bonviewJCCE62027625

RESEARCH ARTICLE

Real-Time Drowsiness Detection Using YOLOv11 in Driver Monitoring Systems

Santhosh Kumar Kadur Lokeshappa[1,*] and **Pravinth Raja Suthakar**[1]

[1] *School of Computer Science and Engineering, Presidency University-Bengaluru, India*

Abstract: One of the main causes of traffic accidents that frequently result in fatalities, serious injuries, and large property losses is driver fatigue. Therefore, enhancing driver safety and preventing such incidents depend heavily on the early, nonintrusive detection of fatigue. A real-time driver monitoring system based on the most recent YOLOv11 framework is presented in this paper. By examining visual indicators like prolonged eye closure and abnormal head movements, the system can identify early indicators of drowsiness. We acquired a purpose-designed dataset that covered a range of driving behaviors and environmental changes in order to improve the framework's flexibility and dependability. The effectiveness of the proposed model was extensively compared with several existing detection methods, and the results indicated that YOLOv11 achieved better accuracy in terms of various evaluation metrics. This improvement is due to its better feature extraction pipeline, attention modules, and computationally efficient architecture, which make it very appealing for real-time applications even inside the vehicle. Overall, the system provides a practical and reliable early warning system for drivers and could greatly reduce the risks associated with fatigue to improve safety in transportation. However, extreme lighting variations and face occlusions may impact the model's performance. To improve robustness, future research will concentrate on expanding the dataset and incorporating multimodal inputs.

Keywords: real-time monitoring, road safety, drowsiness detection, YOLOv11

1. Introduction

Driver fatigue has more and more been regarded as a critical safety problem in modern transportation and is responsible for many traffic accidents around the world. Reaction times and decision-making skills are seriously hampered when the level of alertness is reduced, for example, by long drives, lack of sleep, or irregular duty hours. This leaves drowsiness as one of the most important human factors underlying road accidents. Thus, early identification of such indications of fatigue is essential to reduce the potential risk of accidents and protect road users.

Typically, the existing drowsiness detectors rely on physiological signals such as electroencephalography (EEG), electrooculography, or heart rate variability. While intrusive in nature and requiring direct driver contact, such methods are effective at generating usable results in controlled environments but are impossible to carry out during regular driving conditions.

Vision-based methods, on the other hand, measure driver alertness continuously and noninvasively by using in-car cameras to record behavioral signs like yawning, blinking, extended eyelid closure, or abnormal head orientation.

The accuracy of these monitoring systems has significantly increased with the development of deep learning, especially in computer vision. Both hybrid frameworks and convolutional neural networks (CNNs) have demonstrated efficacy in identifying behavioral and facial fatigue indicators. Finding a balance between generalization to various environmental conditions and computational efficiency is still difficult, though. While highly complex networks might be too resource-intensive for real-time execution inside vehicles, simpler models frequently fail to capture subtle behavioral variations.

The current study presents a drowsiness detection system based on the YOLOv11 architecture in order to overcome these limitations. This model is quick and precise enough to be used in an automobile in real time since it integrates precise classification and quick object recognition into a single pipeline. YOLOv11, which maintains excellent inference efficiency without compromising detection reliability, outperforms several conventional CNN-based techniques.

We created a custom dataset of 1180 labeled images with both awake and drowsy conditions using the Roboflow platform. In this manner, a range of realistic driving scenarios can be used to test the system. The dataset was meticulously divided into training, validation, and testing sets to ensure the accuracy of performance evaluation.

The proposed system creates real-time, annotated visual outputs, classifies the driver's status frame by frame, and uses a temporal threshold-based decision-making mechanism to avoid false positives. Experimental results demonstrate that the YOLOv11-based framework outperforms traditional algorithms such as support vector machine (SVM), Random Forest, and baseline CNN models.

***Corresponding author:** Santhosh Kumar Kadur Lokeshappa, School of Computer Science and Engineering, Presidency University-Bengaluru, India. Email: santhosh.kumarkl@presidencyuniversity.in

The following is a summary of this work's main contributions:

1) Using the YOLOv11 architecture, a nonintrusive, real-time drowsiness detection system is developed.
2) Building a custom dataset with 1180 labeled images of awake and drowsy drivers that has been preprocessed and enhanced for reliable training.
3) Including a temporal decision-making technique to lower frame-wise prediction misclassification.
4) A thorough comparison of the system's accuracy, precision, recall, and F1-score with other baseline algorithms. The upcoming portions of this work are arranged as follows: Section 2 presents the literature review, which emphasizes current approaches. In Section 3, we describe the research method we propose (e.g., data collection, preprocessing, model training, and decision-making). In Section 4, we describe the results and compare them to the baseline approaches. In Section 5, we present a summary of the work and discuss possible pathways for research.

2. Literature Review

2.1. Importance of addressing drowsiness in driving safety

Driver drowsiness remains a serious issue for transportation safety, leading to major traffic accidents worldwide. Drowsiness decreases attention and slows the cognitive reaction time, particularly when driving for long periods of time or at night.

Although capable in controlled contexts, conventional rule-based and handcrafted vision-based systems frequently show poor generalization in real-world scenarios due to variances in illumination, occlusions, and behavioral differences across drivers [1–3]. A shift from handcrafted or single-feature models to adaptive, learning-driven, and multimodal frameworks is necessary to improve system robustness, according to several recent reviews [1, 3, 4].

2.2. Physiological and sensor-based approaches

Since physiological signal monitoring immediately reflects changes in neural and cardiovascular activity brought on by exhaustion, it has long been a primary focus of drowsiness detection research. EEG-based systems have repeatedly shown excellent accuracy among various techniques; for example, detection rates surpassing 90% were reported by Khalil et al. [5], which provides the deep learning approach on top of an embedded system, and Al-Quraishi et al. [6], which gives the sensing technologies used for the driver's state. The electrocardiogram (ECG)-based methods have attained about 80% accuracy [7]; however, their dependability tends to decrease when motion artifacts and car noise are present.

Although they appear to be noninvasive, wearable sensing technologies, including those that use ECG and Photoplethysmography (PPG) readings [8], can be problematic in terms of user comfort and sustained use. In a particularly interesting example, Namburi et al. [9] demonstrated that classification may benefit from integrating EEG data with facial features, as a hybrid classifier achieved 95% accuracy and a 25% relative decline in false alarm rate compared to using either modality alone. While there may be promise with these physiological-based methods, they have issues of sensor intrusion, signal reliability, and practical use in real-world driving scenarios.

2.3. Vehicular dynamics-based methods

Analyzing vehicle telemetry features as secondary indicators of the driver's state of drowsiness including steering angle, lane deviation, and braking force has emerged as a promising line of research. For instance, Lamaazi et al. [10] found that by modeling steering behavior, they achieved an accuracy of 82%, and Pan et al. [11] found that accuracy fell to below 70% when the vehicle was moving on nonuniform surfaces. Liu et al. [12] indicated that by combining telemetry information from steering and acceleration, the accuracy improved to about 89%. Agarkar et al. [13] showed that combining visual input with vehicle dynamics increased overall accuracy to nearly 91%.

According to the review by Egelja and Pavkovic [14], although the reliability of vehicular metrics is significantly affected by external variables such as road surface, driver style, and weather changes, they perform better when combined with visual or physiological components. Thus, even if telemetry-based systems are useful in providing descriptive information, their use of external variables limits their generalizability in a wide range of driving scenarios.

2.4. Vision-based approaches

Vision-based detection has gained traction due to its nonintrusiveness and ease of deployment. After thorough testing, Xu et al. [15] as well as Soukupová and Čech [16] reported accuracies ranging from 75% to 85% for early handcrafted features including blink frequency and PERCLOS. However, their practical efficacy was constrained by inadequate lighting, occlusion, and different driving positions.

Cabot et al. [17] showed that CNN-based models greatly enhanced performance, surpassing 90% accuracy. CNN outperformed handcrafted approaches, as demonstrated by Albadawi et al. [18], who achieved 92.5% in a variety of postures. Essahraui et al. [4] used infrared imaging to solve illumination problems, reporting 92% accuracy when driving at night. To increase resilience, Fu et al. [3] highlighted preprocessing improvements such as denoising and histogram equalization.

These findings demonstrate that vision-based deep learning is more useful than manually created features, even though ambient factors still affect performance.

2.5. Deep learning approaches

Recent research has focused on complex designs such as enhanced CNNs, recurrent neural networks, and attention-based models due to the increasing success of deep learning. After analyzing several CNN variations, Hu et al. [19] found that ResNet-50 could identify eye closure with up to 95% accuracy. In order to model temporal dependencies, Liu et al. [20] investigated CNN–LSTM hybrid frameworks. They achieved 93–94% accuracy; however, their inference performance was only about 8–10 FPS.

Transformer-based methods for drowsiness detection have been introduced more recently. Temporal attention approaches could attain 96–97% accuracy, outperforming CNN–LSTM networks by 2–3% based on studies conducted by Hassan et al. [21] and Zhou et al. [22]. However, the long inference latency of these models—around 200 ms of latency per frame—limited their potential use on vehicles for real-time tasks. Similarly, Venkateswarlu and Ch [23] provided evidence of improved accuracy with Transformer-based designs, but they warned of computational inefficiency relative to CNNs.

This evidence strongly illustrates that although performance is high, any deployment successfully in the real world will depend on low-latency processing in conjunction with capturing, modeling, and utilizing relevant temporal features.

2.6. Multimodal fusion approaches

Combining visual and nonvisual inputs increases the robustness of detection. Yusuf et al. [24] reported that when blink, yawning, and gaze direction were integrated, false alarms decreased by 30%. Phan et al. [25] reported 94% accuracy at just under 95 ms/frame latencies, using a combination of head posture and eye closure.

Audio-visual fusion has also been studied. Ahmed et al. [26] showed that adding speech signals improved accuracy above 95% in simulated tiredness conditions. Namburi et al. [9] demonstrated additional robustness by integrating EEG with facial features. Agarkar et al. [13] and Pan et al. [11] said that more than generalized application, real-time feasibility is impacted as multimodal systems double computational budgets and resources compared to unimodal models.

2.7. YOLO-based approaches for real-time detection

The YOLO family of object detection models is significant to the next evolution of drowsiness detection systems, establishing the optimal balance between accuracy and speed. According to Wang et al. [27], an accuracy and speed of 91.4% and 45 FPS, respectively, were achieved on Jetson TX2. Dhasarathan et al. [28] optimized YOLO into TensorRT, achieving 93.1% accuracy and 65 FPS. How et al. [29] developed a YOLO v8 version classified for night-time driving scenarios, achieving 95.7% accuracy and 52 FPS.

Lakshmi and Vinoth [30] utilized the YOLO v8-FD variant on the NTHU-DDD dataset and found an accuracy of 96.4%. Multi-expert learning modules were incorporated into YOLO v8 by Debsi et al. [31], with accuracy increasing to 97.1%, although GPU acceleration is needed. Chen et al. [32] also developed an attention-based YOLO model that was real-time responsive, achieving 96.8% accuracy while maintaining inference latencies below 50 ms.

Compared to CNN–LSTM architectures (10 FPS) [5, 20] and transformer-based approaches (5 FPS) [21, 22], YOLO-based detectors regularly exceed 45 FPS in inference, making them the most viable and advantageous framework for real-time monitoring of driver drowsiness.

2.8. Datasets and benchmarks

To benchmark models and evaluate performance, using diverse and large-scale datasets is essential. The UTA-RLDD dataset [33] was an important benchmark that reported accuracies between 85% and 90% when using CNN-based architectures. The accuracy went above 93% when using a temporal modeling approach [5]. The same can be said of the NTHU-DDD dataset [34], which is now a common benchmark for fatigue diagnosis using videos. With the use of YOLO-based frameworks, they achieved accuracies of 96% [30].

Majeed et al. [35] also achieved an accuracy of 97.3% with YOLO v8 when it was trained on a customized real-world dataset. This emphasizes the importance of customized datasets. Notably, Fonseca and Ferreira [2] as well as Egelja and Pavkovic [14] emphasized the need for varied datasets to test for robustness and ensure that the model has generalizability outside of the lab to real-world conditions. Dataset diversity means variability in drivers' demographics, lighting, and driving environments.

2.9. Research gaps and insights

The following points are emphasized in this review:

1) Overall, while physiological techniques [5–9] achieve high accuracy, they are invasive and inappropriate for wide application.
2) While vehicular dynamics [10–14] are better when integrated with signals, they are still dependent on the environment.
3) Vision techniques [3, 4, 15–18] face challenges based on illumination and occlusion.
4) Deep learning techniques [5, 19–23] can increase accuracy to 94–97% but often experience latency.
5) Multimodal fusion [9, 13, 24–26] doubles computation costs yet achieves improved robustness.
6) YOLO-based models [27–32] strike the best balance, consistently achieving >45 FPS real-time performance while attaining 95–97% accuracy.

Despite the progress made thus far, accuracy levels of more than 97% would still be a challenge to achieve in real time on edge devices while remaining lightweight. This leads to our proposed YOLOv11-based paradigm.

3. Research Methodology

The drowsiness detection framework proposed is illustrated in Figure 1, which shows the entire system architecture. The first two steps in the sequence of the methodology are data collection and preparation. Model training, classification, decision-making, and visual output come after. As every component is designed to operate in real time, the system is able to monitor the driver's level of attention while driving the vehicle in an effective manner. Furthermore, the modular design ensures the framework is interpretable and is also capable of practical integration.

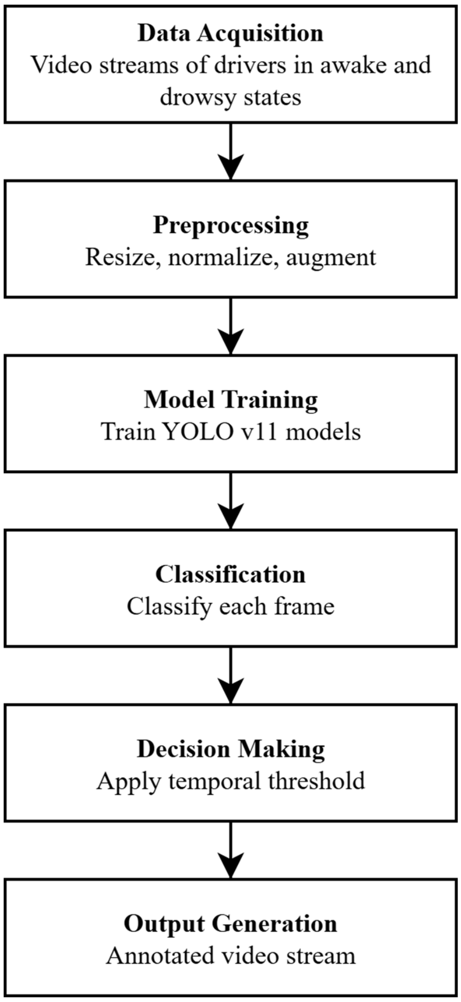
3.1. Data acquisition

Data quality and richness matter significantly for the performance of a vision-based drowsiness monitoring system. To enhance a more accurate detection of drivers' tiredness, a specific dataset was collected from Roboflow.com. The database included 1180 labeled images distributed in two classes (Awake and Drowsy).

Images were captured and chosen in different brightness levels, including artificial light, low light, and light during the daytime, to strengthen the system and maximize its generalization to real-world conditions. And to replicate real-world driver behavior, the head orientation distributions were evenly incorporated.

The Awake class denoted drivers under normal open-eye, upright-eye sitting position and face with alert facial expressions, while the Drowsy class involved head lean, eye closure (with or without sustained eyelid closure), and diversion head position. To address the mild class imbalance between "drowsy" and "awake" samples, data augmentation techniques (e.g., rotation,

Figure 1
System architecture of the proposed real-time drowsiness detection framework using YOLOv11



brightness variation, horizontal flipping) were applied to the minority class.

The use of a curated task-oriented dataset rather than a public benchmark contributed to the fact that the designed model can effectively generalize across various conditions. This tailored database could contribute to establish a solid foundation for developing a reliable real-world monitoring system that can work in real driving situations.

3.2. Data preprocessing

Before model development, the dataset was preprocessed in a standard way to be compatible with the YOLOv11 framework. All images were preprocessed into 640 × 640 pixels, conforming to the input requirement of the model and equalizing scales among samples. This scaling operation not only reduced the variability in the dimension of the image, but it was also able to exploit GPU resources to efficiently learn visual patterns.

All other preprocessing was handled by the tools on the Roboflow platform besides resizing. The pixel intensity distributions across all datasets were normalized, and data augmentations such as random rotations, brightness/contrast adjustment, and horizontal flips were employed to introduce additional diversity into the data. These are also indicative of real-world driving and help the model generalize.

This preprocessing procedure made the dataset computationally light and rich in information, which, in turn, enhanced the robustness and generalization power of the devised drowsiness detection framework.

3.3. Model training

The dataset was split into three separate subsets for training and testing: 70% for training, 20% for validation, and 10% for testing. This allocation by strata guaranteed the presence, at the level of the subsets, of both classes (Awake and Drowsy) in similar proportions, generally avoiding bias and improving model validity.

Training was performed with a batch size of 16, thus managing to keep a good trade between the amount of used GPU memory and the stability of gradient updates. We started the learning rate with a value of 0.001, which is reduced gradually by the cosine decay scheduler during training for better convergence over training iterations. A momentum of 0.937 was used in the Stochastic Gradient Descent optimizer for faster convergence, while a weight decay factor of 0.0005 was used to prevent overfitting.

The classifier was trained for up to 100 epochs via cross-entropy loss, which distinguishes between "Awake" and "Drowsy" states by reducing misclassifications. In order to avoid unnecessary computations and prevent overfitting, early stopping was always used: when the loss on the validation set did not improve for 20 epochs, we quit fitting.

With this configuration, the YOLOv11-based model maintained computational efficiency and generalization capacity while learning representative visual cues of driver drowsiness through important visual features and context information, such as facial landmarks. The trained model clearly distinguishes between awake and drowsy driving states and makes accurate predictions on sample inputs, as seen in Figure 2. The annotated outputs show how well the model handles a variety of dataset circumstances.

3.4. Classification

The classification phase determines whether the driver is in an Awake or Drowsy state using the trained YOLOv11-based classifier. Each input video frame I_t at the time of t is resized to 640 × 640 pixels and passed through the classification network. The model outputs class probabilities using the softmax function, defined as

$$P(c \mid I_t) = \frac{e^{z_c}}{\Sigma_{j=1}^{C} e^{z_j}},\ c \in \{Awake,\ Drowsy\}$$

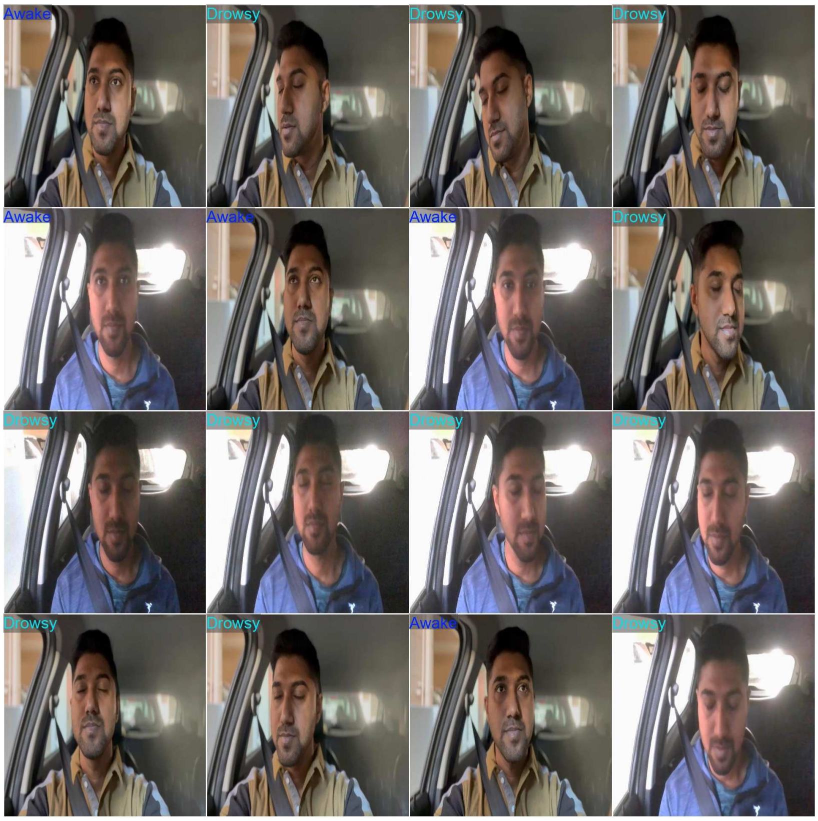
where z_c represents the logit score for class c, and C is the total number of classes (in this case, 2). The predicted class label for the frame is assigned as

$$\hat{y}_t = \arg\max_{c} P(c \mid I_t)$$

To mitigate transient misclassifications caused by eye blinks, illumination variations, or partial occlusions, a temporal smoothing strategy is employed. A counter D is maintained to track consecutive Drowsy predictions. For each frame:

$$D = \begin{cases} D + 1, & if\ \hat{y}_t = Drowsy \\ 0, & if\ \hat{y}_t = Awake \end{cases}$$

Figure 2
Illustrative batch predictions from the YOLOv11 drowsiness detection framework



As illustrated in Figure 3, the proposed YOLOv11-based framework performs real-time frame-wise classification by detecting the driver's face and labelling the state as either Awake or Drowsy.

Figure 3
Frame-wise classification using the proposed YOLOv11 model on real-time video streams for drowsiness detection

(a)
(b)

3.5. Decision-making

The decision layer converts per-frame class predictions into a stable, time-consistent driver state. Blinks, head twists, motion blur, and light transients are common causes of isolated errors, even though the classifier individually generates an Awake/Drowsy label for each frame. We use a temporal validation method based on a consecutive-frames threshold, along with confidence gating and failure handling, to reduce such noise. Next, a thresholding rule determines the final choice label L for the driver state:

$$L = \begin{cases} Drowsy, & if\ D > T \\ Awake, & otherwise \end{cases}$$

where $T =$ 25 frames in this study. Both frame-rate features and human blink patterns shown in drowsiness detection experiments were used to objectively calculate the value. In order to ensure robustness against false alarms, the system only labels the driver as tired when the model predicts consecutive drowsy states over a period of more than 25 frames. This hybrid method enhances both sensitivity to real drowsiness and specificity against false positives by fusing frame-level probability estimation with sequence-level decision rules.

3.6. Output generation

The visual output generation, the last phase of the proposed drowsiness detection system, offers a user-friendly and instantaneous feedback method for tracking the driver's level of attentiveness. The classification result is superimposed on the video frames to graphically convey whether the driver is "Awake" or "Drowsy," as determined by the decision-making module. Bounding boxes around identified face regions are used by the system, and the color of the box indicates the alertness state—green for "Awake" and red for "Drowsy". Additionally, as shown in Figure 4, the matching label is shown above the bounding box to ensure that the end user can easily see the classification result.

Even in real-time driving situations, the output is immediately comprehensible due to the combination of textual labels and bounding box color coding. To produce the final visual output that evidently differentiates between Awake and Drowsy states, the system integrates face detection, state classification, and real-time description. Figure 5 illustrates the workflow of the real-time drowsiness detection system, while Table 1 provides the complete algorithm.

Beyond that, the processed video stream is saved as an output file for offline studies, providing utility for data set expansion, performance evaluation, and research. The system also furnishes two active feedback levels via text-based and visual signs, providing better reliability and applicability for the monitoring solution.

Figure 4
Final output generation of the proposed YOLOv11-based drowsiness detection system

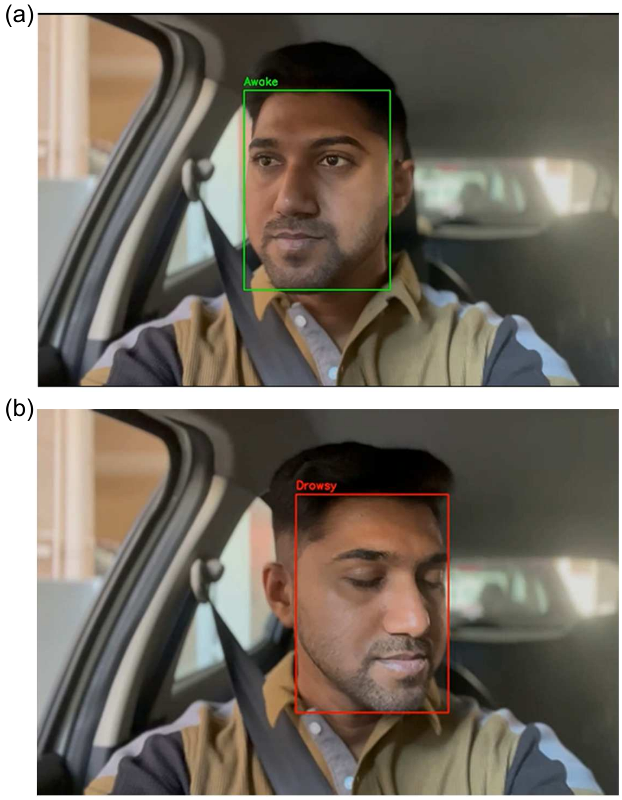


Figure 5
Workflow of the proposed drowsiness system

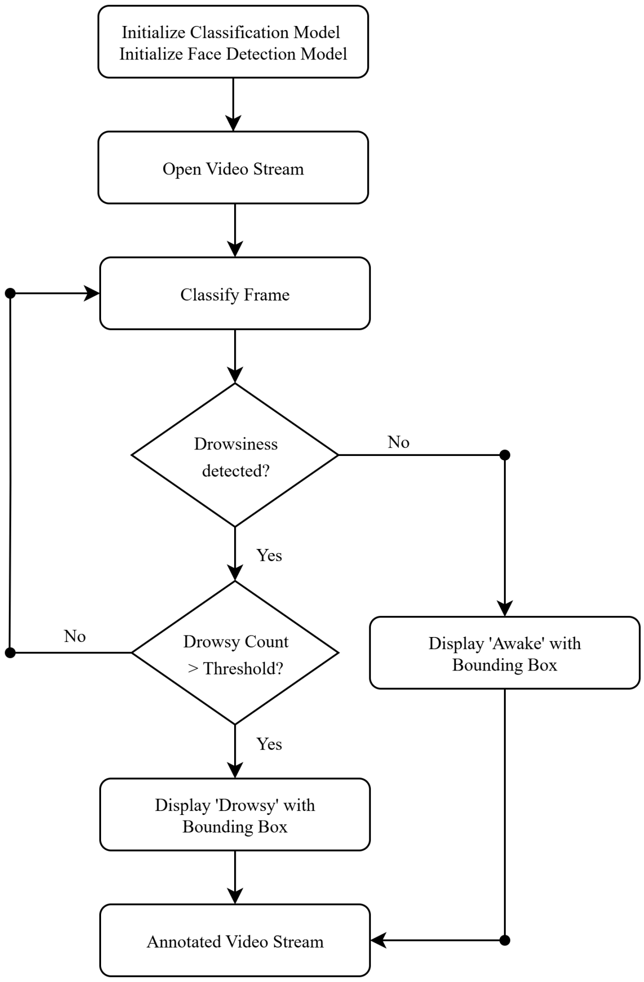


Table 1
Algorithm for the proposed drowsiness system

Algorithm: Proposed YOLOv11-based Drowsiness Detection Framework
Input: a. Live video stream containing driver's face b. Pretrained YOLOv11 classification model (M_{cls}) c. Face detection model (M_{det}) d. Threshold value T for consecutive drowsy frame count **Output:** Annotated real-time video stream with bounding box with classification label (*Awake* or *Drowsy*)
1. Initialize models Load the classification model (M_{cls}) Load the face detection model (M_{det}) **2. Start input stream** Open the live video stream **3. Frame acquisition and classification** Capture the current frame F Apply M_{cls} to classify F as either Awake or Drowsy

(*Continued*)

Table 1
(*Continued*)

Algorithm: Proposed YOLOv11-based Drowsiness Detection Framework
4. Update drowsiness counter If the frame is classified as *Drowsy*, increment the counter Otherwise, reset the counter to *Zero* **5. Decision logic** If the drowsiness counter ≥ threshold T → Set the final label as *Drowsy* Else: → Set the final label as *Awake* **6. Face detection and annotation** Use M_{det} to detect the face region in frame F For each detected face: Draw a bounding box Annotate with the final label *Awake or Drowsy* **7. Output display** Display the annotated frame in real time **8. Repeat** Continue Steps 3–7 until the video stream is terminated **9. End process** Close the video stream

4. Result and Discussion

The custom dataset created through Roboflow was used in a number of studies to assess the efficacy of the proposed drowsiness detection methodology. A total of 1180 labeled images representing a variety of driving states, including both awake and drowsy situations, under various lighting and head-position settings, were included in the dataset. The dataset was split into 70% for training, 20% for validation, and 10% for testing in order to maximize training efficiency and provide an objective evaluation. The Ultralytics YOLOv11 framework was used to develop the entire system in Python. The T4 GPU runtime environment, which offered sufficient computational capabilities for both model training and real-time inference testing, was used to conduct experimental trials on the Google Colab platform.

4.1. Performance metrics

Widely recognized performance metrics like accuracy, precision, recall, and the F1-score were used to assess the categorization framework's efficacy. While precision and recall provide insights on the trade-off between missed detections and false alarms, accuracy quantifies the total ratio of accurate predictions. In situations like drowsiness detection, when reducing false alerts is essential to preserving user confidence, the F1-score—which integrates precision and recall into a single harmonic mean—is particularly helpful.

4.1.1. Accuracy

Accuracy represents the proportion of correctly classified instances (both Awake and Drowsy) relative to the total number of samples. It is expressed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP denotes true positives (drowsy instances correctly identified), TN true negatives (awake instances correctly identified), FP false positives (awake instances misclassified as drowsy), and FN false negatives (drowsy instances misclassified as awake).

While accuracy gives an overall measure of system performance, it can be misleading in cases of class imbalance.

4.1.2. Precision

Precision quantifies the proportion of correctly predicted drowsy frames out of all frames predicted as drowsy:

$$Precision = \frac{TP}{TP + FP}$$

A high precision value indicates that the system produces fewer false alarms, which is particularly important in drowsiness detection, as excessive false alerts can cause driver annoyance and reduce trust in the system.

4.1.3. Recall

Recall measures the proportion of actual drowsy frames that were correctly detected by the system:

$$Recall = \frac{TP}{TP + FN}$$

High recall ensures that the system successfully identifies most drowsy events, minimizing the risk of missed detections that could lead to hazardous driving situations.

4.1.4. F1-score

The F1-score is the harmonic mean of precision and recall, offering a balanced measure when there is a trade-off between false alarms and missed detections:

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

This metric is particularly useful in safety-critical applications such as drowsiness detection, where both precision (avoiding unnecessary alarms) and recall (detecting actual drowsy cases) must be optimized simultaneously.

4.2. Confusion matrix

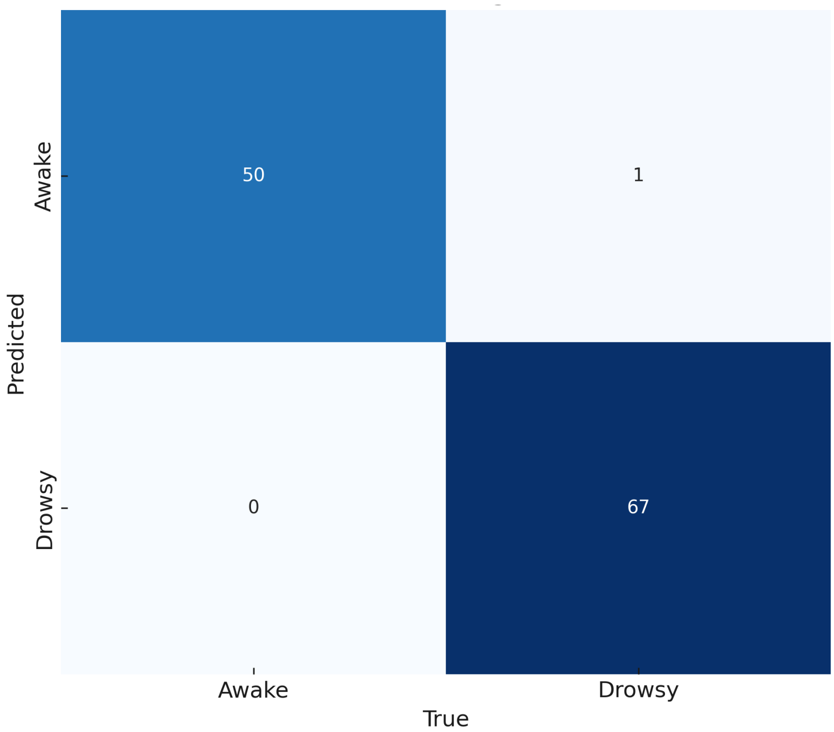
As seen in Figure 6, a confusion matrix was used to further evaluate the efficacy of the proposed YOLOv11-based drowsiness detection methodology. The distribution of samples that were correctly and wrongly identified for both driving states is shown in the matrix. The model correctly identified 50 instances of the Awake class and 67 instances of the Drowsy class from the evaluation set.

There were very few misclassifications: 0 drowsy frames were mistakenly labeled as awake, and 1 awake frame was improperly predicted as drowsy.

Given that it exhibits a high sensitivity to drowsiness and an adequate specificity in the alert state, this finding demonstrates the model's remarkable discriminative potential.

The YOLOv11 framework's robustness is confirmed by the low rates of false positives and false negatives. The system's applicability for real-time driver monitoring, which necessitates prompt and precise answers and has a minimal chance of misclassification, is justified by its balanced performance across several classes.

Figure 6
Confusion matrix for YOLOv11-based drowsiness detection



4.3. Comparative analysis with existing algorithms

The performance evaluation results presented in Table 2 and Figure 7 demonstrate the efficiency of the proposed YOLOv11-based drowsiness detection scheme with the conventional and deep learning structures. All the baseline models were trained and tested using the same dataset employed for the proposed YOLOv11 model.

However, accuracies of traditional machine learning methods, including SVM and Random Forest classifiers, were less than 90%, which meant they could only ensure moderate prediction reliability. Their reliance on hand-designed features made them less robust to variations in driver pose, lighting, and occlusions.

Furthermore, these techniques all demonstrated quite low recall, signifying that the systems were making many incorrect predictions of drowsiness that compromise safety in real-world applications. Deep learning models, such as a customized CNN and a hybrid model with CNN layers combined with Long Short-Term Memory (LSTM), were shown to markedly outperform conventional models. The CNN achieved an accuracy of more than 91%, while the LSTM-CNN hybrid could get 93.8%, which demonstrates the advantage of temporal sequence learning. But due to computational requirements, real-time performance was limited even though they had better generalization; that is, they had even lower FPS compared with YOLOv11.

However, in this study, the proposed YOLOv11 significantly outperformed all baselines that resulted in 99.1% accuracy, 98.5% precision, 100% recall, and 99.2% F1-score by real-time inference at 25 FPS. The large recall of drowsy samples guarantees reliable detection, avoiding false-negative alarms, and the performance of high precision avoids false positive alarms. This balance makes the framework both practically deployable and highly accurate.

Reasonably, by comparing the performance of YOLOv11-based drowsiness detection system, we can conclude that the YOLOv11 framework can achieve a better trade-off between performance accuracy and computational efficiency and is a comfortable and dependable model for a driver monitoring system in real time.

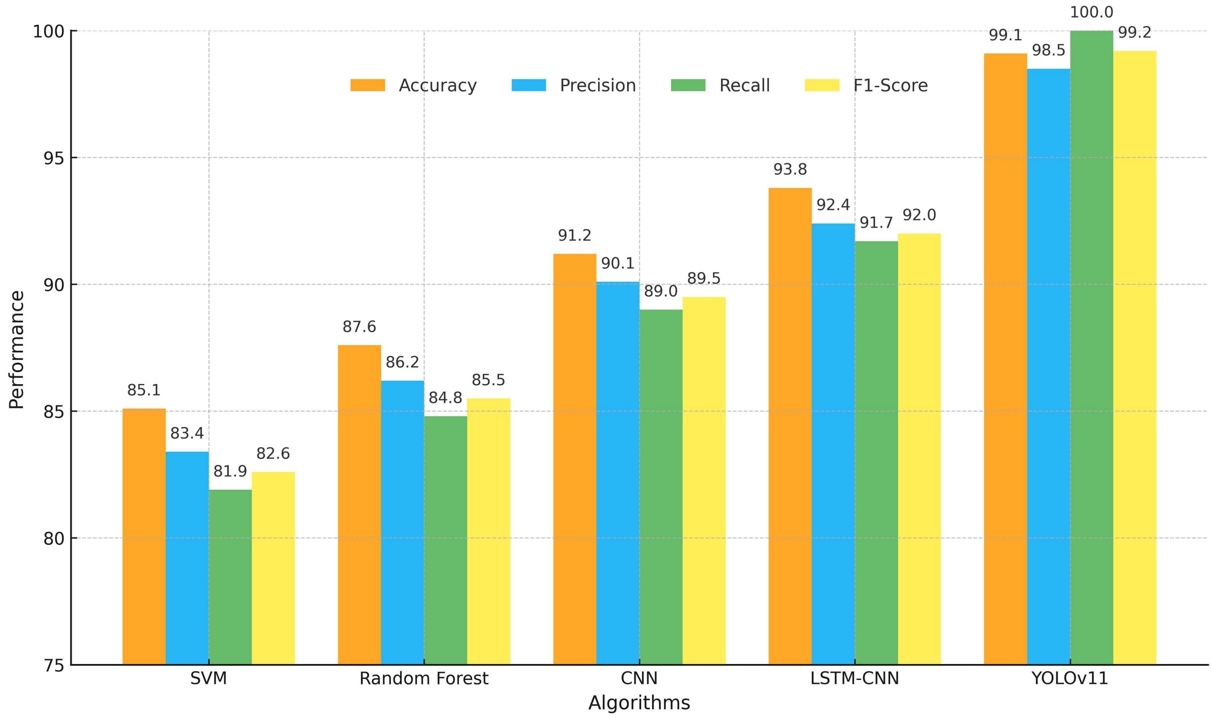
4.4. Qualitative observations

In addition to numerical results, visual inspection of the processed video streams further validated the performance of the system in driver state distinction. The addition of bounding boxes with color coding (green for awake, red for drowsy) and text labels provided intuitive feedback that can be easily understood

Table 2
Performance metrics for different algorithms

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
SVM	85.1	83.4	81.9	82.6
Random Forest	87.6	86.2	84.8	85.5
CNN	91.2	90.1	89.0	89.5
LSTM-CNN hybrid	93.8	92.4	91.7	92.0
Proposed YOLOv11	99.1	98.5	100	99.2

Figure 7
Comparative performance of drowsiness detection algorithms



by the users. Importantly, the decision-making module was robust to noise, meaning that neither normal eye blinks nor short gaze shifts were falsely identified as signs of fatigue. These findings further corroborate the robustness of the system in routine driving scenarios.

4.5. Discussion

The results show that the proposed method can achieve high detection performance and temporal consistency and can run in real time. The main reasons for this performance are the introduction of a customized dataset, a well-organized preprocessing pipeline, and a comprehensively devised training strategy.

Nevertheless, some limitations were identified. When the driver's face is extremely occluded (e.g., with big sunglasses or facing away from the camera), the detection is slightly unreliable at the moment. To overcome these challenges, further developments of this work could consider other modalities, for example, physiological monitoring, and control patterns of the vehicle to increase robustness and reliability of the system.

5. Conclusion and Future Work

In this paper, we present a real-time system for driver drowsiness detection that is designed based on the YOLOv11 framework and fuses facial region localization, frame-level classification, and information from the time domain to provide reliable alertness monitoring of a driver.

A customized dataset containing 1180 labeled images, augmented and preprocessed to a uniform resolution of 640 × 640 pixels, was employed to train and validate the system. Experimental evaluation demonstrated that the proposed approach achieved superior performance compared to traditional machine learning models and deep learning baselines, with accuracy, precision, recall, and F1-scores all exceeding 98%.

The system successfully generated annotated video outputs in real time, highlighting the driver's state as either Awake or Drowsy with color-coded bounding boxes. These findings indicate that the framework based on YOLOv11 is computationally efficient and can be practically applied for intelligent driver-assistance systems. While the YOLOv11-based drowsiness detection system delivered high accuracy and reliable real-time performance, numerous future opportunities for improvement are expected.

An initial option could involve expansion to multimodal sensing for more effective drowsiness detection, which would supplement the visual analysis modality with physiological signals (EEG, heart rate variability) or motion analysis of behavioral characteristics (steering wheel motion).

Another option is to improve robustness to challenging conditions (low lighting, partially occluded face, masks or sunglasses as accessories). The custom dataset may also impose limitations for demographic diversity (ethnicity, gender, facial structure), which may cause bias when generalizing to broader demographics.

As future work, the proposed YOLOv11 driver monitoring system can be experimented on the well-known benchmark datasets such as UTA-RLDD [33] and NTHU-DDD [34] to prove the strength of the approach using a wide variety of driving scenarios.

To enable on-board driver monitoring without relying on high-performance cloud resources, edge optimization and deployment on embedded platforms, such as automotive-grade microcontrollers or GPUs, should also be investigated.

Last but not least, extensive pilot tests with a variety of driver demographics and actual driving circumstances will offer more information on the system's dependability, usability, and long-term effectiveness in lowering fatigue-related accidents.

Recommendations

Even though the proposed system displayed good real-time capabilities and attained high accuracy, there are still a number of areas that might be improved.

First, adding more real-world driving scenarios to the dataset such as different lighting, occlusions, and facial orientations would improve generalization even further.

Second, adding multimodal data such as physiological signals, audio cues, or steering patterns could improve robustness in complicated contexts where visual clues might not be enough.

Third, wider usage in low-power automotive systems would be made possible by improving the model for lightweight edge deployment on embedded devices.

Finally, conducting large-scale field testing with diverse driver populations will provide stronger validation of the system's applicability in real-world traffic conditions.

Acknowledgment

The authors are grateful to Presidency University, Bengaluru, India, for providing the support and infrastructure for the research.

Ethical Statement

This study involving human participants was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Institutional Ethics Committee of Presidency University, Bengaluru (approval reference no.: PSCS/EAI/2026/01).

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in Roboflow website at https://universe.roboflow.com/santhosh-kumar-k-l/drowsiness-detection-koqqc.

Author Contribution Statement

Santhosh Kumar Kadur Lokeshappa: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Pravinth Raja Suthakar:** Resources, Writing – review & editing, Visualization, Supervision.

References

[1] Albadawi, Y., Takruri, M., & Awad, M. (2022). A review of recent developments in driver drowsiness detection systems. *Sensors*, *22*(5), 2069. https://doi.org/10.3390/s22052069

[2] Fonseca, T., & Ferreira, S. (2025). Drowsiness detection in drivers: A systematic review of deep learning-based models. *Applied Sciences*, *15*(16), 9018. https://doi.org/10.3390/app15169018

[3] Fu, B., Boutros, F., Lin, C.-T., & Damer, N. (2024). A survey on drowsiness detection: Modern applications and methods. *IEEE Transactions on Intelligent Vehicles*, *9*(11), 7279–7300. https://doi.org/10.1109/TIV.2024.3395889

[4] Essahraui, S., Lamaakal, I., El Hamly, I., Maleh, Y., Ouahbi, I., El Makkaoui, K., . . . , & Abd El-Latif, A. A. (2025). Real-time driver drowsiness detection using facial analysis and machine learning techniques. *Sensors*, *25*(3), 812. https://doi.org/10.3390/s25030812

[5] Khalil, H. A., Hammad, S. A., Abd El, Munim, E. H., & Maged, S. A. (2025). Low-cost driver monitoring system using deep learning. *IEEE Access*, *13*, 14151–14164. https://doi.org/10.1109/ACCESS.2025.3530296

[6] Al-Quraishi, M. S., Ali, S. S. A., Al-Qurishi, M., Tang, T. B., & Elferik, S. (2024). Technologies for detecting and monitoring drivers' states: A systematic review. *Heliyon*, *10*(20), e39592. https://doi.org/10.1016/j.heliyon.2024.e39592

[7] Sahayadhas, S., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: A review. *Sensors*, *12*(12), 16937–16953. https://doi.org/10.3390/s121216937

[8] Shen, Q., Zhao, S., Zhang, R., & Zhang, B. (2020). Robust two-stream multi-features network for driver drowsiness detection. In *Proceedings of the 2020 2nd International Conference on Robotics, Intelligent Control and Artificial Intelligence*, 271–277. https://doi.org/10.1145/3438872.3439093

[9] Namburi, A., Sitpasert, P., & Duang-onnam, W. (2024). A CNN-LSTM approach for accurate drowsiness and distraction detection in drivers. *ICIC Express Letters*, *18*(9), 907–917. https://doi.org/10.24507/icicel.18.09.907

[10] Lamaazi, H., Alqassab, A., Fadul, R. A., & Mizouni, R. (2023). Smart edge-based driver drowsiness detection in mobile crowdsourcing. *IEEE Access*, *11*, 21863–21872. https://doi.org/10.1109/ACCESS.2023.3250834

[11] Pan, H., Logan, D. B., Stephens, A. N., Payre, W., Wang, Y., Peng, Z., . . . , & Koppel, S. (2025). Exploring the effect of driver drowsiness on takeover performance during automated driving: An updated literature review. *Accident Analysis & Prevention*, *216*, 108023. https://doi.org/10.1016/j.aap.2025.108023

[12] Liu, G., Wu, K., Lan, W., & Wu, Y. (2025). YOLO-FDCL: Improved YOLOv8 for driver fatigue detection in complex lighting conditions. *Sensors*, *25*(15), 4832. https://doi.org/10.3390/s25154832

[13] Agarkar, A. S., Gandhiraj, R., & Panda, M. K. (2023). Driver drowsiness detection and warning using facial features and hand gestures. In *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies*, 1–6. https://doi.org/10.1109/ViTECoN58111.2023.10157233

[14] Egelja, M., & Pavkovic, B. (2024). A review of the approaches for enhancing driver safety through camera-based monitoring. In *Zooming Innovation in Consumer Technologies Conference*, 36–41. https://doi.org/10.1109/ZINC61849.2024.10579113

[15] Xu, C., Huang, W., Liu, J., & Li, L. (2025). Detecting driver drowsiness using hybrid facial features and ensemble learning. *Information*, *16*(4), 294. https://doi.org/10.3390/info16040294

[16] Soukupová, T., & Čech, J. (2016). Real-time eye blink detection using facial landmarks. In *21st Computer Vision Winter Workshop*.

[17] Cabot, N., Lamouchi, D., Yaddaden, Y., & Cherif, R. (2024). Drowsiness detection through yawning and eye blinking models using convolutional neural networks and transfer learning. In *2024 Sixth International Conference on Intelligent*

Computing in Data Sciences, 1–8. https://doi.org/10.1109/ICDS62089.2024.10756434

[18] Albadawi, Y., AlRedhaei, A., & Takruri, M. (2023). Real-time machine learning-based driver drowsiness detection using visual features. *Journal of Imaging*, *9*(5), 91. https://doi.org/10.3390/jimaging9050091

[19] Hu, Z., Wu, D., Nie, F., & Wang, R. (2021). Generalization bottleneck in deep metric learning. *Information Sciences*, *581*, 249–261. https://doi.org/10.1016/j.ins.2021.09.023

[20] Liu, M.-Z., Xu, X., Hu, J., & Jiang, Q.-N. (2022). Real-time detection of driver fatigue based on CNN-LSTM. *IET Image Processing*, *16*(2), 576–595. https://doi.org/10.1049/ipr2.12373

[21] Hassan, O. F., Ibrahim, A. F., Gomaa, A., Makhlouf, M. A., & Hafiz, B. (2025). Real-time driver drowsiness detection using transformer architectures: A novel deep learning approach. *Scientific Reports*, *15*(1), 17493. https://doi.org/10.1038/s41598-025-02111-x

[22] Zhou, F., Alsaid, A., Blommer, M., Curry, R., Swaminathan, R., Kochhar, D., . . . , & Tijerina, L. (2022). Predicting driver fatigue in monotonous automated driving with explanation using GPBoost and SHAP. *International Journal of Human–Computer Interaction*, *38*(8), 719–729. https://doi.org/10.1080/10447318.2021.1965774

[23] Venkateswarlu, M., & Ch, V. R. R. (2024). DrowsyDetectNet: Driver drowsiness detection using lightweight CNN with limited training data. *IEEE Access*, *12*, 110476–110491. https://doi.org/10.1109/ACCESS.2024.3440585

[24] Yusuf, L., Hamada, M., Hassan, M., & Kakudi, H. (2023). Enhanced driver drowsiness detection model using multi-level features fusion and a long-short-term recurrent neural network. *Engineering Proceedings*, *56*(1), 338. https://doi.org/10.3390/ASEC2023-15537

[25] Phan, A.-C., Trieu, T.-N., & Phan, T.-C. (2023). Driver drowsiness detection and smart alerting using deep learning and IoT. *Internet of Things*, *22*, 100705. https://doi.org/10.1016/j.iot.2023.100705

[26] Ahmed, M. I. B., Alabdulkarem, H., Alomair, F., Aldossary, D., Alahmari, M., Alhumaidan, M., . . . , & Zaman, G. (2023). A deep-learning approach to driver drowsiness detection. *Safety*, *9*(3), 65. https://doi.org/10.3390/safety9030065

[27] Wang, Z., Yao, K., & Guo, F. (2023). Driver attention detection based on improved YOLOv5. *Applied Sciences*, *13*(11), 6645. https://doi.org/10.3390/app13116645

[28] Dhasarathan, C., Gnanasekaran, S., Pattanayak, A., Kumar, G., Vig, K., Narain, V., . . . , & Garg, S. (2025). Tensor RT optimized driver drowsiness detection system using edge device. *Ain Shams Engineering Journal*, *16*(10), 103620. https://doi.org/10.1016/j.asej.2025.103620

[29] How, H. W., Mahamad, A. K., Saon, S., Othman, M., Kadir, H. A., & Elmunsyah, H. (2024). Development of drowsiness & drunk detection using YOLOv8. In *2024 International Conference on Future Technologies for Smart Society*, 6–11. https://doi.org/10.1109/ICFTSS61109.2024.10691374

[30] Lakshmi Devi, S., & Vinoth, A. (2025). Real-time driver fatigue detection and classification using YOLOv8-FD: A precision-driven approach for enhanced road safety. *Indian Journal of Science and Technology*, *18*(26), 2111–2123. https://doi.org/10.17485/IJST/v18i26.732

[31] Debsi, A., Ling, G., Al-Mahbashi, M., Al-Soswa, M., & Abdullah, A. (2024). Driver distraction and fatigue detection in images using ME-YOLOv8. *IET Intelligent Transport Systems*, *18*(10), 1910–1930. https://doi.org/10.1049/itr2.12560

[32] Chen, C., Liu, X., Zhou, M., Li, Z., Du, Z., & Lin, Y. (2025). Lightweight and real-time driver fatigue detection based on MG-YOLOv8 with facial multi-feature fusion. *Journal of Imaging*, *11*(11), 385. https://doi.org/10.3390/jimaging11110385

[33] Ghoddoosian, R., Galib, M., & Athitsos, V. (2019). A realistic dataset and baseline temporal model for early drowsiness detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 178–187. https://doi.org/10.1109/CVPRW.2019.00027

[34] Weng, C.-H., Lai, Y.-H., & Lai, S.-H. (2017). Driver drowsiness detection via a hierarchical temporal deep belief network. In *Computer Vision – ACCV 2016 Workshops* (pp. 117–133). https://doi.org/10.1007/978-3-319-54526-4_9

[35] Majeed, F., Shafique, U., Safran, M., Alfarhood, S., & Ashraf, I. (2023). Detection of drowsiness among drivers using novel deep convolutional neural network model. *Sensors*, *23*(21), 8741. https://doi.org/10.3390/s23218741

How to Cite: Lokeshappa, S. K. K., & Suthakar, P. R. (2026). Real-Time Drowsiness Detection Using YOLOv11 in Driver Monitoring Systems. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE62027625