

A Real-Time Driver Drowsiness Detection Using Computer Vision and Machine Learning

P.Swathy^{1,*}, Bajjuri Meghana², Nalla shreevansh³, Brahmadevara SriVarsha⁴ and Rajanala Siri⁵

¹⁻⁵ Department of Computer Science and Engineering (AI and ML), Kakatiya Institute of Technology and Science, Warangal, Telangana -506015

*Corresponding Author: P.Swathy, Email: sp.csm@kitsw.ac.in

Received: 07/04/2026; Revised: 25/04/2026; Accepted: 16/05/2026; Published: 30/06/2026.

DOI: <https://doi.org/10.69996/fmep.2026011>

Abstract: Driver drowsiness is one of the main factors that lead to road accidents, thereby causing serious injuries and deaths. Driver monitoring systems that are traditionally adopted include physical or physiological measures like EEG, ECG, or wearables, which are cumbersome, expensive, and not suitable for practical use. In order to overcome such disadvantages, a new non-intrusive approach for driver drowsiness detection based on their vision is presented. In the proposed approach, Haar cascade classifiers are utilized, resulting in effective facial feature detection like face, eye, and mouth, based on live webcam feed. All detected regions are further processed using a YOLO-based classification model for classifying and deciding the driver's status like eye-open, eye-closed, yawn, and combined conditions of eye closed with a yawn. On the basis of these classifications, the level of alertness of the driver is effectively determined. In case of drowsiness, a three-level alert system is implemented in the system, viz. buzzer sound, voice, and video recording. If drowsiness persists, the recorded video is saved and sent via email.

Keywords: Driver Drowsiness Detection, Augmentation, YOLOv-8, Haar Cascade, Real-Time Monitoring.

1. Introduction

Driver drowsiness is considered one of the major contributing factors to road accidents worldwide. A fatigued driver has a reduced attention span, slower reaction times, and short microsleep episodes that may even go unnoticed by the driver [1]. These minor-looking lapses may end in disastrous or fatal accidents in the blink of an eye. This is why early detection of drowsiness has become critical in modern road safety systems. Traditional solutions for drowsiness detection generally rely either on vehicle-based measures, such as steering patterns, lane deviation, or pedal movement, or on physiological signals including EEG and heart rate [2]. While these methods have some potential, many of them are costly, invasive, and involve substantial hardware requirements, making their deployment on a large scale problematic. The use of vision-based driver monitoring systems has attracted major attention to overcome these limitations. Such systems offer non-intrusive, cost-effective, and highly scalable solutions for real-time fatigue detection by using just a simple camera and deep learning algorithms [3-5]. They analyze facial cues such as eye closure, blinking rate, yawning, and head orientation to accurately assess the driver's alertness level. In this project, we propose a real-time driver drowsiness detection system powered with YOLOv8-a state-of-the-art object detection model renowned for its speed and accuracy [6]. It will be used along with Haar Cascade classifiers, which will detect the key facial regions around the eyes, face, and mouth efficiently. These regions, once identified, are then classified for various driver states: Open Eyes, Closed Eyes,



Yawn, and No-Yawn. Based on these, the system will automatically trigger a multilevel alert mechanism, including buzzer alarms, voice warnings, video evidence recording, and email notifications to designated members [7-10]. Such progressive alerts will offer immediate feedback to the driver while capturing the necessary evidence in critical situations for safety monitoring [11-12]. It should also work reliably in real driving conditions-for example, different lighting conditions, partial occlusions, and natural head movements[13-15]. It will thus combine deep learning methods with practical alerting mechanisms to contribute to improved road safety by offering a robust framework for the prevention of accidents due to driver drowsiness.

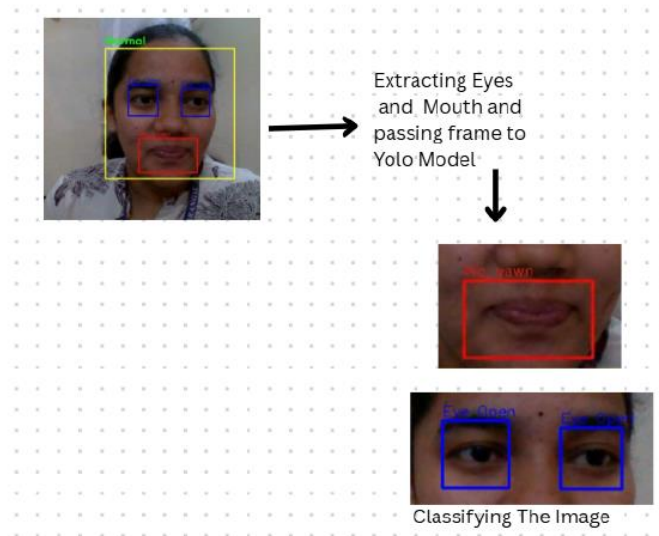


Figure 1: Overview of Real-time driver drowsiness detection

2. Related work

The related work describes the progress and success in detecting driver drowsiness in real-time through the use of computer vision and deep learning. The earlier works were based on traditional feature-derived methods using Haar cascades computation and some machine learning classifiers, which verified that blinking, eyelid closure and the frequency of yawning could be the main indicators of drowsiness. Recent works are more concentrated on deep learning architectures such as CNNs, vision transformers, hybrid models, which have better performances and robustness to various illumination and driving environments. A number of whitepapers have pointed out challenges such as needing to be real-time, model lightness, as well as multi-feature fusion for robust drowsiness detection while driving. To summarize, all the results established in the literature, clearly demonstrate the transition from hand-crafted features detection to advanced deep learning approaches, and underline some gaps and potentialities for the realization of efficient real-time frameworks such as the framework developed in this study.

In 2025, Essahraui et al. proposed a real-time driver drowsiness detection algorithm with facial landmark extraction, eye aspect ratio (EAR) and mouth opening ratio ratio (MOR) based on the integration of machine learning classifiers. The system has proved to work very well under varying illumination and different face poses. However, there are still challenges for their model when the driver wears glasses or there are partial occlusions, which decreases the accuracy of landmark detection. While the system performs well in controlled settings, its accuracy is compromised by rapid head movements or when the face is not facing the camera directly.

Hassan et al. 2025, proposed a new drowsiness detection approach that utilized the power

of Vision Transformer architectures. Unlike the traditional CNN-based methods, this model used self-attention to capture temporal patterns, such as the frequency of blinking and the duration of yawning. This provided a highly improved classification accuracy for drowsiness in long-duration driving. This work highlighted the advantages of transformer models for analyzing sequential changes of the face but pointed out that ViT models required high computational resources, hence limiting their use on low-power embedded systems.

Srinithi et al. (2025) proposed a CNN-based drowsiness detection model associated with Haar Cascade for face detection. In their approach, a CNN was employed to classify open and closed eye states based on features extracted from the cropped region-of-interest. This approach yielded highly promising results under strongly controlled conditions. However, when drivers tilted their heads or when their images were blurred because of motion, the accuracy of eye detection seriously degraded. The use of a standard Haar cascade limited the robustness of face detection under poor illumination conditions.

In the work of Jarndal et al. (2024), Vision Transformers were employed in a more advanced vehicle safety system for driver drowsiness detection in real time. Their model ensured high robustness under low-light conditions, glare, and partial occlusion of the driver's face while outperforming several CNN architectures. Further, the system demonstrated good accuracy over variation of camera angle and driver posture. However, the authors reported higher inference time compared to lighter models, hence less suitable for real-time deployment with devices with a limited GPU capability.

Jiao (2023) investigated the correlation of physiological signals with driver fatigue through the use of heart rate variability and electrodermal activity. The study classified alert and fatigued states using a Gradient Boosting classifier, which achieved an accuracy of 88.7%. Despite the high reliability of physiological signals, the system was unable to be applied in practical, real-time automotive use due to the need for specialized sensors. The study concluded that measures based on vision were superior in non-intrusive monitoring.

Chand (2022) proposed a drowsiness detection system based on the analysis of emotional state using CNN models. The model tried to identify emotional fatigue, which generally coincides with drowsiness by evaluating facial expressions and micro-emotions. The system achieved an accuracy of 73.23% in establishing the viability of emotion-driven fatigue detection. However, when drivers showed neutral facial expressions or subtle emotional cues, it detracted from the real-world application of the model.

In the year 2021, Zhang et al., proposed a driver fatigue monitoring system based on PERCLOS measurement, where the percentage of eye closure over time measured helps to determine drowsiness in drivers. The system distinguishes the eye region using image processing techniques, followed by the detection of fatigue based on the closure of the eyes. The system effectively detects fatigue in a controlled environment, but the system is extremely sensitive to changes in illumination, face occlusions, as well as head movements, thereby reducing the accuracy of the system.

In the study by Kaur et al., 2021, the researchers designed a yawning detection model using deep learning technology. A convolution neural network has been used to train the model on different facial image datasets. The model can learn significant features by itself to detect yawning behavior. This model provides better accuracy in yawning detection but needs a large dataset for training purposes. In addition, it consumes more computational resources.

Pan et al., 2020, proposed a YOLO scheme for driver monitoring by employing object detection algorithms. The video frames were processed, and facial features were detected for

determining the alertness of the driver. The model was able to achieve fast detection and was appropriate for real-time applications. However, the method focuses only on frame-by-frame analysis and may generate false alarms due to the lack of time modeling.

Abtahi et al., 2014, suggested a driver drowsiness monitoring system based on yawning detection through facial landmark extraction and mouth opening. In this system, yawning detection leads to fatigue detection based on the change in the aspect ratio of the mouth. In spite of effectively detecting yawning, this method is limited, as it only takes into account the mouth and omits other possible characteristics used for fatigue, including eye closure and blinks.

Orosco et al., 2014, proposed a blink rate analysis method to check driver drowsiness using eye movement monitoring. In this method, driver fatigue is assessed in terms of eye blink count and eye closure time. Although the method offers a non-invasive method to check driver alertness levels, the eye blink rate shows significant variability among different individuals due to external factors.

Park et al., 2016, presented a deep learning approach for the construction of a driver drowsiness detecting system based on the learning of feature representations from facial characteristics by employing various structures of neural networks. This method enhances the process of detecting drowsiness through the extraction of features from the obtained images. Nevertheless, the process requires significant computational resources and complex training of the model.

3. Dataset Description

To Prepare the dataset the System The driver drowsiness detection system herewith proposed this is modeled after a structured approach encompassing the preparedness of the dataset model training, detection application in the real-time generation of the alarms. The major components of this proposed system can be defined as follows:

a. Preparation of the dataset (Dataset Preparation):

In reference to the Driver Drowsiness Detection dataset, a detailed dataset is established. The dataset contains images of faces for different driver states: open eyes, closed eyes, yawning, and non-yawning. The dataset is employed as a necessary device for an accurate deep learning based classification. For more reliability and generalization, various lighting conditions, head position, facial orientation and expression are considered. The typical dataset like DDD provides a true and perfect training and detecting for the driver drowsiness.

b. Data Augmentation:

In order to improve the precision as well as robustness of the proposed model, data augmentation is employed. The transformations applied are rotation to account for head tilt; scaling to simulate the different distances of the camera and horizontal flipping to simulate the left-right motion of the head. Also brightness is compensated for low light and overexposed driving scenes. These augmentation methods increase the diversity of the datasets substantially and the resulting models generalize better in real driving with reduced overfitting, and thus better performance.

c. Splitting the Dataset:

The augmented data set is organized systematically into three subsets to guarantee the training and evaluation of the system is done with certainty. 80% of the dataset is allocated for training, which is to learn the feature representations for training the YOLOv8 model. 10% is used for a validation set for hyperparameter optimization and for monitoring the models performance to stop training prior to overfitting. The other 10% is the testing set, which is applied to assess the predictive accuracy and generalization capacity of the model on unseen data.

This allows the proposed system to be fully evaluated before actual system implementation.

d. Model Training (YOLOv8):

In order to carry out accurate classification of a driver's state, the YOLOv8 model is trained using the annotated dataset. The objective of the trained model is to effectively classify the four main classes, which include eye open, eye closed, yawn, and no-yawn. However, in order to ensure accurate classification, a number of optimization procedures are carried out. These include scheduling the learning rates, normalizing the batch, and monitoring the loss, all of which ensure that accurate classification is achieved by the model in a fast inference framework.

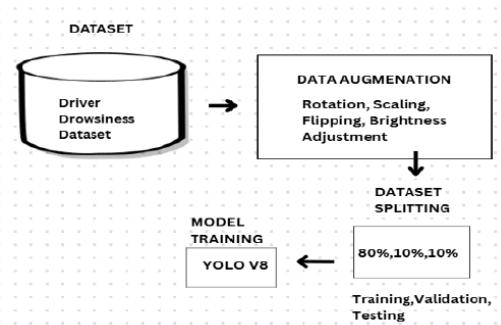


Figure 2: Data Setup and Model Training

e. Real-Time Face Detection:

Once the training stage is completed, a real-time driver monitoring is conducted with the live video being taken directly through a webcam by the proposed system. Each video frame received is analyzed with Haar Cascade classifiers to quickly find and identify facial features such as the face, two eyes, and the mouth. Based on the low computational complexity and good real-time performance of Haar cascade classifier, system can extract the desired ROIs in each frame with very small amount of computation and high speed. These localized ROIs are powerful in removing redundant background and making sure only relevant facial features are searched. This two-stage architecture achieves the best trade-off between accuracy and computational complexity and is general enough to be applied for real-time driver drowsiness detection.

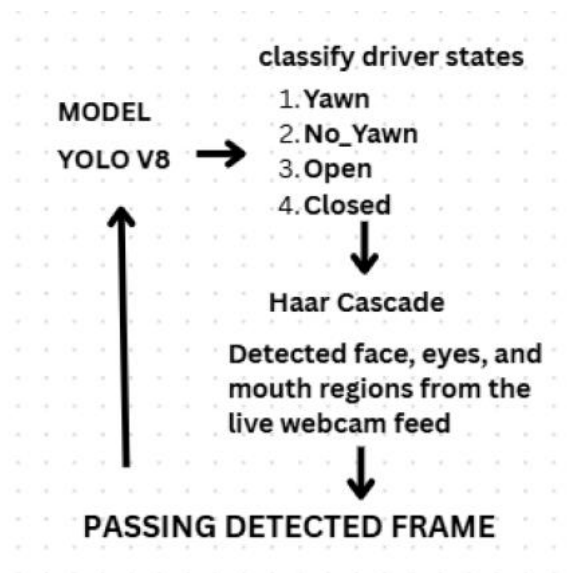


Figure 3: Real Time Detection of Driver

f. Continuous Monitoring and Alert Mechanism:

Instead, it finds the driver's drowsiness by considering the prediction pattern over time rather than single-frame decisions. The continuous algorithm monitors the duration of eye closure and yawning in terms of seconds continuously to find reliable discrimination between normal blinking and actual drowsiness. Further, based on the duration and severity of drowsiness detected, a multilevel alert mechanism is activated to improve driver safety. Initial warnings for brief indications of fatigue are provided, supported by buzzer alerts for sustained eye closure in high-risk conditions. Moreover, voice warnings are triggered in immediate dangerous conditions to draw the driver's attention back. In critical levels of drowsiness persistence for larger durations, evidence-based email notifications are automatically generated. This consigns progressive alert policy so intervention receives both on time and chance drowsy behavior accident can be minimized.

g. Alert System and Video Proof:

A multi-level warning procedure in line with the length and severity of detected driver sleepiness will be applied by the presented system to guarantee that the driver receives sufficient attention in good time. A short warning fatigue signal is given for short eye closes, while buzzers alarms sounds in the event of long eye closures. For very high risk situations, voice alarm is activated to maximize recover the driver's attention. In the event of critical drowsiness detection for a long time, it automatically sends e-mail to the pre-registered e-mail addresses There are also instances where the system can generate a 5 seconds evidence video which is stored locally and included to the alert email in case of extreme situations. Such video-driven documentation is a method that may be used to inform and appropriately intervene by supervisors and safety officers. Overall, it will be progressive alerts with recorded evidence that will enhance accountability on the roads.

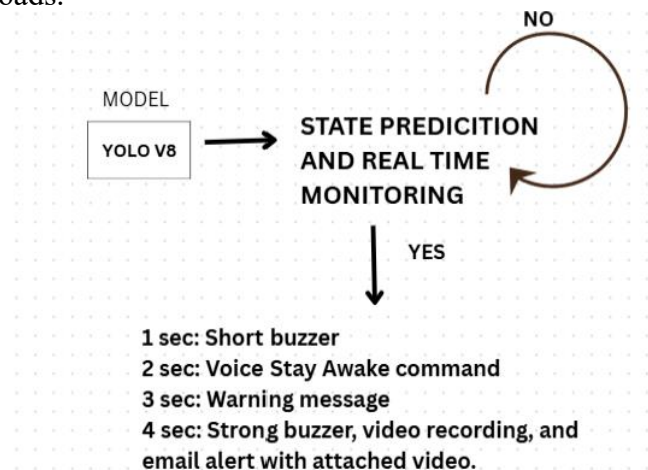


Figure 4: Continuous Monitoring and Real Time Updates

4. Experimental Setup

A. YOLOV8 Architecture

YOLOv8 is an object detection and classification architecture described as single-stage and anchor-free, highly accurate, and consistent with real-time requirements. The architecture is composed of three main components:

1. Backbone:

The backbone network performs the task of feature extraction from the input images. YOLOv8 utilizes the CSP-based (Cross Stage Partial) model with convolutional layers for

efficient feature extraction from various facial regions: eyes and mouth.

2. Neck:

The neck employs Feature Pyramid Network (FPN) along with Path Aggregation Network (PAN) for multi-scale feature fusion. This ensures successful detection of even minor facial changes such as closed eyes and yawning.

3. Head:

Classification and confidence prediction are carried out by the detection head in an anchor-free fashion. In this work, classification of the extracted facial region is performed into four classes representing different driver states such as eye open, eye closed, yawn, and no yawn. This makes the classifier computationally simpler, enabling faster inference. This is because the model's lightweight design makes it suitable for real-time driver monitoring.

B. Algorithm and Workflow Input:

Video stream from webcam Driver alert status and evidence of recording Initializing the webcam and the trained YOLOv8 model with the Haar cascade classifiers. For each captured frame Apply Haar cascade classifiers to detect the face region.

1. To locate the eyes and mouth within the recognized face, the regions of interest (ROI) can be taken.
2. The ROIs, after extraction, will be preprocessed by resizing and normalizing them according to the YOLOv8 requirements.
3. Input the preprocessed ROIs into the pre-trained model YOLOv8 for driver state classification.
4. Each of the frames will be classified under four categories: eye-open, eye-closed, yawn, and no-yawn.
5. The prediction outcomes need to be saved with timestamp information and be observed as a function of time to obtain eye closure/yawn period.
6. Comparison of duration measured against certain thresholds for quantifying the level of driver drowsiness.
7. Generate multi-level alerts depending on the level of drowsiness severity.
8. Warns of transient fatigue laterInitially shows the warning for transient fatigue. Activate buzzer on long eye closure. Voice Alert for high risk states.
9. For critical drowsiness, send an e-mail alert. A five-second video clip is recorded for the extreme cases. Then the video is locally stored and sent via emails as an attachment. This will continue until the video stream is stopped.

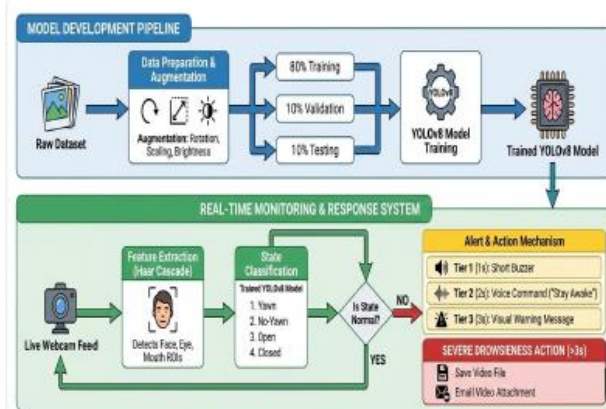


Figure 5: System Architecture

5. Results and Discussion

In the execution stage, the trained YOLOv8 model was deployed along with facial features extraction in real-time and alerting system. The YOLOv8 model was trained on an augmented dataset based on four major driver states: Open Eyes, Closed Eyes, Yawn, and No-yawn. Rotational, scale, span of brightness and flipping augmentations have been done to make it consistent with real-world scenarios.

A. Model training

The model was trained over 50 epochs with a batch size of 16 and the use of GPU accelerator for faster computation and convergence. The powerful YOLOv8 architecture definitely also facilitated better and efficient learning of minutiae that are essential for differentiation of eye and mouth states. In proposed method to enhance quality of input for YOLO model was combined with Haar Cascade classifiers to detect face, eyes and mouth in each frame taken from the webcam. These were then cropped and passed into the yolov model for classification. This mixed method reduces wasteful computations, and enhances the accuracy of classifications. During co-execution, the following operations executed in order:

1. The system reliably detected the face and accurately located the eye/ mouth region under most illumination and pose conditions. Trustworthy Detection: The face was correctly detected, and the eye and mouth region were localized in the majority of the illumination and pose conditions.
2. State Classification: YOLOv8 yielded stable predictions for all four classes and could capture the differences among normal blinking, extended eye closure, yawning, and none-yawning.
3. Activation of The Alerting Mechanism: On the basis of monitoring the system continuously, tiered alert would be triggered: buzzer alarms and voice warnings would be triggered at the initial stage and final serious warning would be issued when the drowsiness was detected for an extended period of 16.5 minutes.
4. Automated Video Logging: When severe drowsiness over the threshold level occurred, the system captured 5-second video clippings and dispatched them as attachments with automated email notifications, demonstrating the viability of emergency response function. In summary, the proposed solution was successfully evaluated and found to be effective, stable, and applicable to real-world situations.

B. Results

The proposed method was further validated through extensive real-time experiments and the results were satisfactory on the given test dataset. The trained YOLOv8 model was proven to be the model with the best performance among all the evaluation metrics. The results proved that the model can still finely distinguish the driver's states between drowsy and non-drowsy in some extreme cases such as the driver's eyes only slightly closed or there are subtle symptoms of drowsiness. Besides classification rate, the system also achieves seamless real-time operation with about on a commodity laptop. Such a processing rate is indeed sufficient for real-time surveillance, as this would allow the system to generate alarms without any delay. In addition, the combined alarm and alert system was effective. Depending on how long a sleepiness is detected, the procedure to alert the driver includes to simple visual warnings to loud buzzer sound along with voice messages. In the event that drowsiness is sustained or critical, a state of emergency protocol is initiated. This protocol captures a 5 seconds video and sends it through automated mail alerts which increase the trustworthiness and usability of the system. On the whole, this success conclusion of these functions ensures that the proposed driver drowsiness

detection system is not only stable for real time implementation but also responsive & reliable for real time implementation.



Figure 6: Validation Batch SET-1



Figure 7: Validation Batch SET-2

During model validation, ground-truth labels will be used to generate sample outputs by YOLOv8 in the form of predictions on the validation images, which have not been observed before.

val_batch0_labels refers to the true labels annotated of the DDD dataset. And at each image there are manual labels of driver states: eye open, eye closed, yawn, no-yawn. These annotations are considered as reference ground truth for the comparison of model performances.

pred: val_batch0 The model prediction on the same images but from the validation set and not training. Predictions here are the class detected by the YOLOv8 model using learned feature representations and not any prior knowledge.

Visually, with respect to frames, one can observe that the predicted labels are more or less consistent with the ground-truth annotation in both figures while 弦 be changed by varying illumination, rotation of the face, and different facial expression. This demonstrates that our suggested hybrid framework (Haar cascade ROI extraction + YOLOv8 classification training on DDD dataset) is effective and efficient in the wild.

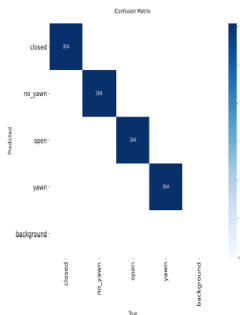


Figure 8: Confusion Matrix

The confusion matrix resulting from this trained YOLOv8 driver state classification model indicates excellent classification performance in all the classes, as shown below:

Each driver state category, such as eye closed, eye open, yawn, and no yawn, showed a correct prediction result of 304 with no signs of off-diagonal values existing in the results. This implies that the model was successful in learning discriminative features for the classification of the eye and mouth state based on the DDD dataset, including the application of the relevant data augmentation methods. The results also proved the relevance and efficiency of the integration of Haar cascade for the extraction of ROI in the implementation of the YOLOv8 classification algorithm, saving from processing irrelevant features for classification purposes.

The absence of classification errors highlights good class separability and good

generalization of the model on the test data set. These results validate that the suggested system can correctly separate between normal blinking, prolonging the closure of eyes, and yawning actions with high precision to detect and warn drowsiness with high accuracy.

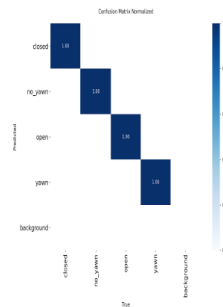


Figure 9: Confusion Matrix Normalized

The normalized confusion matrix shows the performance of a classifier in terms of percentages rather than the actual number of instances. Each measure is standardized to lie between 0 and 1. This shows the percentage of instances correctly and incorrectly classified per class. From the obtained matrix, all the diagonal values are 1.00, which implies that all the samples in their respective classes were classified correctly at 100%. The lack of non-zero values along the off-diagonal implies that there are no classification errors in any of the four drive states: eye closed, open, yawn, or no yawn. It is evident that the results yield 100% precision and recall for all the categories, thus supporting the efficiency of the proposed hybrid model. The model seamlessly combines Haar Cascade-based ROI detection with a YOLOv8 classifier (trained on the DDD dataset) to provide precise feature localization and classification. In addition, the normalized confusion matrix is also strong class separability, showing that the model well discriminates the different driver states. This enhanced separability can be regarded as an indication of a better generalization. All in all, these results demonstrate the robustness and reliability of the proposed approach, which validates its merit to be used for real time driver drowsiness detection applications.

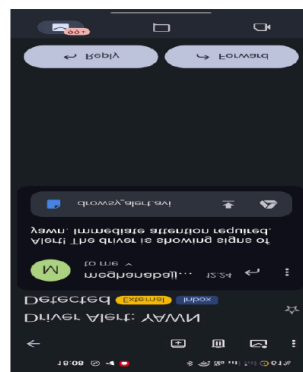


Figure 6: Mail Alert

The above image represents the automatic email notification generated by the driver monitoring system. The subject of the email, “Driver Alert: YAWN Detected,” shows that the model has recognized a yawn event for driver drowsiness. The email message states: “Alert! The driver is showing signs of Yawn. Immediate Attention Required.” Additionally, the system also includes a video file (drowsy_alert.avi) as evidence of the identified event. This provides an example of the real-time alert mechanism provided by the hybrid system, in which results from the Haar Cascade + YOLOv8 classifier (with the DDD dataset) are communicated to the end

user.

```
PS D:\Real_Time_000> python evaluation.py
curves_results: {}
fitness: 1.0
keys: ['metrics/accuracy_top1', 'metrics/accuracy_top5']
results_dict: {'metrics/accuracy_top1': 1.0, 'metrics/accuracy_top5': 1.0, 'fitness': 1.0}
save_dir: mindweath\m\Real_Time_000\res(classify\val10)
speed: {'preprocess': 0.26543061041923, 'inference': 2.698280511254731, 'loss': 0.00575481236461154, 'postprocess': 0.0002761618109863}
task: 'classify'
top1: 1.0
top5: 1.0
Top Accuracy : 1.0
PS D:\Real_Time_000>
```

Figure 7: Yolo Accuracy

```
C:\Users\varsh\AppData\Local\Programs\Python\Python312\Lib\site-packages\keras\sources\trainers\data_adapters\py_dataset_adapter.py:122: UserWarning: Your 'PyDataset' class should call 'super().__init__(**kwargs)' in its constructor. '**kwargs' can include 'workers', 'use_multiprocessing', 'max_queue_size'. Do not pass these arguments to 'fit()', as they will be ignored.
  self._warn_if_super_not_called()
114/114 — 31s 264ms/step - accuracy: 0.9856 - loss: 0.0364
Test Loss = 0.0500
Test Accuracy = 0.9825
PS D:\CNN>
```

Figure 8: CNN Accuracy

Table 1: Performance Table

Model	Task Role in System	Test Accuracy	Key Strength
YOLOv8	Real Time Eye - Mouth State Classification	1.0	Fast Detection
CNN	Image Classification of eye and mouth states	0.98	Stable learning

Considering the above performance comparison, it can be noted that the hybrid model, i.e., YOLOv8-based model, has achieved better classification accuracy when compared to the performance of the CNN model. Also, it can be observed that the YOLOv8 model has achieved perfect top-1 accuracy, i.e., 1.0, during evaluation, indicating better potential for distinguishing between driver states like blinking, yawning, and prolonged closure of eyes due to the usage of Haar Cascade classifiers.

Moreover, the CNN model showed promising results as the test accuracy was 98.25%, and the loss was minimal. This proved the effectiveness of the CNN model as a classifier. But, when comparing the results of the CNN model with the YOLOv8 model, the YOLOv8 showed better results as a real-time classifier with improved classification accuracy. Finally, the experimental results validate the reliability of the hybrid method as a solution for the proposed problem.

6. Conclusion

In this article, the authors proposed their real-time driver drowsiness detection system, which leverages the detection of the face, eyes, and mouth through the hybridization of the Haar Cascade classifier and the YOLOv8 deep learning classifier. With their proposed model, they efficiently detected the driver conditions of normal blinking, prolonged closure of eyes, yawning, and non-yawning states. The results of the experiment were reliable and real-time; the YOLOv8 performed better than the baseline CNN model with relatively higher accuracy. With the incorporation of alert mechanisms such as buzzer alerts, voice alerts, and the sending of email alerts with 5-second video clips, the system is safe for drivers. The system performed well as it was found to be stable, lightweight, and applicable in real-world situations. It was concluded that the proposed framework was useful in minimizing computational costs while maximizing detection precision.

Acknowledgment: Not Applicable.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

- [1] S. Essahraoui, I. Lamaakal, I. El Hamly, Y. Maleh, I. Ouahbi et al., "Real-time driver drowsiness detection using facial analysis and machine learning techniques," *Sensors*, vol. 25, no. 3, pp. 812, 2025.

- [2] O.F. Hassan, A. F. Ibrahim, A. Gomaa, M. A. Makhoul and B. Hafiz, "Real-time driver drowsiness detection using transformer architectures," *Scientific Reports*, vol. 15, pp. 17493, 2025.
- [3] A. Jarndal, H. Tawfik, A.I. Siam, I. Alsyoud and A. Cheaitou, "A vision transformers-based system for enhanced driver drowsiness detection," *IEEE Access*, 2024.
- [4] C. Srinithi, G. Lathaselvi and A. Swetha, "Predicting driver drowsiness using CNN and real-time facial detection," *Proc. ICCRTEE*, pp. 1–6, 2025.
- [5] Y. Jiao, "Driver fatigue detection using heart rate variability and electrodermal activity," *Sensors*, vol. 23, 2023.
- [6] H.V. Chand, "Driver drowsiness detection system using emotion analysis," *International Journal of Intelligent Systems*, 2022.
- [7] X. Zhang, "PERCLOS-based driver fatigue monitoring system," *IEEE Trans. Intelligent Vehicles*, 2021.
- [8] G. Kaur, "Deep learning-based driver yawning detection," *Pattern Recognition Letters*, 2021.
- [9] Y. Pan, "YOLO-based driver monitoring system," *Journal of Real-Time Image Processing*, 2020.
- [10] M. Abtahi, S. Hariri and S. Shirmohammadi, "Driver drowsiness monitoring based on yawning detection," *IEEE Instrumentation & Measurement Magazine*, 2014.
- [11] A. V. Orosco, J. A. Gonzalez, and E. Laciari, "Blink rate analysis for driver drowsiness detection," *International Journal of Pattern Recognition and Artificial Intelligence*, 2014.
- [12] S. Park, F. Pan, S. Kang and C. D. Yoo, "Driver drowsiness detection system based on feature representation learning using various deep networks," *Asian Conference on Computer Vision*, 2016.
- [13] Swara Snehit Patil, "Artificial Intelligence: A Way to Promote Innovation", *JSIHS*, vol. 2, no. 1, pp. 1–5, Mar. 2024.
- [14] Kavin Ramakrishnan and Velmurugan S, "Machine Learning and Sensor-Based Approaches for Driver Fatigue and Drowsiness Detection: A Mini Review", *JCAI*, vol. 4, no. 1, pp. 1–10, Feb. 2026.
- [15] Godfrey Wandwi and Gwenda Wandwi, "Deep Learning Architectures for Early Detection of Diabetic Retinopathy in Retinal Image Analysis", *JCAI*, vol. 3, no. 5, pp. 68–75, Oct. 2025.