DRIVER DROWSINESS DETECTION AND ALERT SYSTEM DEVELOPMENT USING YOLOv8

Nguyen Manh Trung[1,*], Tran Hung Cuong[1]

DOI: https://doi.org/10.57001/huih5804.2025.423

ABSTRACT

Drowsy driving is a common cause of serious traffic accidents, especially during long journeys or under fatigue conditions. This study develops a real-time drowsiness detection and warning system for drivers to enhance traffic safety. The method uses the YOLOv8 deep learning model to recognize characteristic behaviors such as prolonged eye closure, yawning, and head nodding through image data from cameras. The system requires no physical contact, has fast processing capabilities, and operates stably under various lighting conditions and viewing angles. The data used for training and testing the model is the Drowsiness Detection XSRIZ dataset. Experimental results show that the model achieves high performance with a mAP@0.5 score of 98.6%, ensuring reliability in drowsiness state recognition. The system can effectively detect and warn of driver drowsiness in real-time, helping drivers to promptly adjust their behavior or stop to rest when necessary.

Keywords: *Driver Drowsiness Detection, Drowsiness, DDD, Driver Safety.*

[1]School of Information and Communications Technology, Hanoi University of Industry, Vietnam
[*]Email: nmt2811203@gmail.com
Received: 12/8/2025
Revised: 25/9/2025
Accepted: 28/11/2025

1. INTRODUCTION

An Road traffic safety remains a pressing global concern amidst the rapid development of transportation systems. Among the various factors contributing to traffic accidents, driver drowsiness has emerged as a significant threat to human life and socio-economic stability. Accidents related to drowsiness often occur suddenly, leaving drivers unable to respond promptly to emergencies, resulting in severe casualties and substantial property damage. In Vietnam, according to the Ministry of Public Security, out of a total of 21,880 road traffic accidents in 2023, approximately 72 cases (accounting for 0.33%) were attributed to driver fatigue and drowsiness [1]. A 2023 study published in Scientific Reports revealed that 11.5% of traffic accident patients in Vietnam reported feeling tired and drowsy prior to driving [2]. Globally, the U.S. National Safety Council [3] reported that around 13% of drivers admitted to falling asleep at the wheel at least once per month, and 4% had caused accidents due to drowsiness. The most vulnerable times for drowsy driving are between 2:00 - 6:00a.m. and 2:00 - 4:00p.m.

Recognizing the critical nature of this issue, leading automobile manufacturers such as Mercedes-Benz, Tesla, and Volkswagen have developed driver drowsiness detection technologies. However, these systems are predominantly available in high-end vehicles with premium pricing, making them inaccessible to the general public. Initial research approaches utilized biological sensors such as EEG and EOG to measure physiological signals, but these methods are invasive and uncomfortable for drivers. Subsequently, traditional image processing techniques, such as Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), were adopted. While these are non-invasive, they often lack accuracy and robustness under varying environmental conditions.

The emergence of deep learning has opened a new chapter in this field. Deep learning models possess the capability to automatically learn and extract complex features from image data, overcoming the limitations of traditional methods. Among various deep learning architectures, YOLO (You Only Look Once) stands out for its unique advantage of detecting objects in a single processing step, achieving an optimal balance between speed and accuracy. Unlike conventional multi-stage pipelines, YOLO can simultaneously detect and classify multiple objects within a single frame, making it particularly suitable for real-time applications such as driver state monitoring. YOLOv8, the latest version in the YOLO family, demonstrates significant improvements in performance and accuracy over previous versions. With its optimized architecture and fast processing capabilities, YOLOv8 is an ideal choice for detecting drowsiness-related cues such as prolonged eye closure, yawning, and head nodding in real-time.

This study develops a comprehensive driver drowsiness detection and alert system based on YOLOv8, capable of accurately identifying drowsiness signs using non-contact camera input. The system is designed to operate in real time, providing timely audio alerts to raise driver awareness of fatigue and drowsiness. Upon detecting signs such as prolonged eye closure, repetitive yawning, or head nodding, the system automatically triggers an alert, advising the driver to take a break or hand over control of the vehicle, thereby reducing the risk of drowsiness-related accidents. This research focuses on building and evaluating the effectiveness of the YOLOv8-based drowsiness detection system, aiming to deliver a cost-effective and widely deployable solution for real-world applications.

2. RELATED WORK

Driver drowsiness detection is a critical research domain in enhancing traffic safety, in which object detection techniques play a central role. These techniques can be broadly categorized into two groups: traditional and deep learning-based approaches. Traditional methods, such as Histogram of Oriented Gradients (HOG), Haar-like features, and Canny edge detection, typically rely on handcrafted feature extraction followed by classification using models such as Support Vector Machines (SVM), Random Forests, or simpler algorithms like Naive Bayes. These methods are relatively easy to implement but often require complex preprocessing steps and generally achieve only 70 - 85% accuracy under real-world conditions. Their performance is highly susceptible to environmental factors such as low lighting, unstable camera angles, and occlusion, which limits their suitability for real-time monitoring applications.

The advancement of deep learning has introduced a new paradigm with significantly improved performance. Deep learning-based object detection techniques are generally divided into two categories: two-stage and one-stage detectors. Two-stage algorithms such as R-CNN and Fast R-CNN [4] detect objects through a two-step process involving region proposal generation followed by classification. These models achieve high accuracy for complex tasks but are computationally intensive and exhibit latency, making them less suitable for real-time applications. In contrast, one-stage detectors, including You Only Look Once (YOLO) [5] and Single Shot MultiBox Detector (SSD) [6], combine detection and classification into a single processing step, offering significantly faster inference speeds and greater feasibility for deployment on embedded systems. Due to their balance of accuracy and speed, one-stage detectors are increasingly favored in real-time surveillance applications. YOLO, in particular, has demonstrated high efficiency in processing image data streams rapidly, making it ideal for driver monitoring in complex vehicular environments.

Numerous studies have employed both traditional and deep learning techniques to detect driver drowsiness by analyzing behavioral features such as blinking, yawning, and head posture. For instance, the study by Phan Anh Cang et al. [9] proposed a hybrid system combining VGG16, InceptionV3, and DenseNet with Long Short-Term Memory (LSTM) networks, leveraging transfer learning to identify fatigue-related signs such as head tilt, eye blinking, and yawning. This system outperformed facial feature–based methods [9] by exploiting DenseNet's feature extraction capabilities and LSTM's temporal modeling. However, the system still struggled to maintain performance under complex lighting conditions. Another study by Yaman Albadawi et al. focused on visual cues, employing Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and head pose estimation in conjunction with classifiers such as Random Forest, Sequential Neural Network, and Linear SVM [14]. This approach achieved an impressive 99% accuracy on the National Tsing Hua University Driver Drowsiness Detection Dataset, but further validation is required under more complex and realistic driving scenarios.

Similarly, M.I. Basheer Ahmed et al. implemented a Convolutional Neural Network (CNN) integrated with VGG16 to analyze eye movement and facial expressions, achieving an accuracy of 95.1% on a dataset of 2,900 images depicting eye states and yawning [10]. However, this method also faced challenges in low-light or suboptimal camera angle conditions. Another approach by Yaman Albadawi and colleagues integrated visual features such as EAR, MAR, and head pose with physiological signals like EEG or heart rate [13]. Using machine learning methods including SVM, Random Forest, and CNN, the system achieved an average accuracy of 93.33% using the PERCLOS (Percentage of

Eye Closure) metric. Despite promising results, the requirement for multi-modal data integration increased system complexity and implementation costs.

Studies utilizing YOLO have shown remarkable potential in drowsiness detection due to its speed and accuracy. For example, Balaji et al. employed YOLOv8 in combination with CNN to detect drowsiness indicators, achieving an accuracy of 85.6% - a promising result, though slightly limited, potentially due to inadequate training or lack of dataset diversity [11]. Another study by Ghanta Sai Krishna et al. combined YOLOv5 with Vision Transformers (ViT) for facial detection and drowsiness classification, achieving an accuracy of 97.4% on a custom dataset of 39 participants under various lighting conditions [12]. However, this approach required large volumes of annotated data and further optimization for deployment on resource-constrained devices. Compared to two-stage methods such as Faster R-CNN, which are computationally expensive and unsuitable for real-time applications, YOLO offers superior performance for driver monitoring. Earlier YOLO versions such as YOLOv3, YOLOv4, and YOLOv5 were progressively improved to balance speed and accuracy, yet continued to struggle in complex environments involving variable lighting or partial occlusions.

Building upon these foundations, YOLOv8 [7] introduces significant enhancements, particularly in terms of robustness and adaptability to diverse environmental conditions. Without radically altering its core architecture, YOLOv8 is optimized for improved performance in real-world scenarios, including low-light settings, dynamic head poses, and partial facial occlusions common challenges in drowsiness detection. These improvements enable YOLOv8 to extract more nuanced facial features of drivers while maintaining high processing speed and accuracy, making it well-suited for real-time monitoring systems. This study proposes leveraging YOLOv8 to detect fatigue-related symptoms using holistic facial image inputs, rather than relying solely on isolated features like EAR or MAR. The goal is to build a robust, efficient, and practical driver monitoring solution with high potential for real-world deployment, thereby contributing to enhanced traffic safety.

3. PROPOSED METHODOLOGY

3.1. Introduction to YOLOv8

❖ ***Architecture***

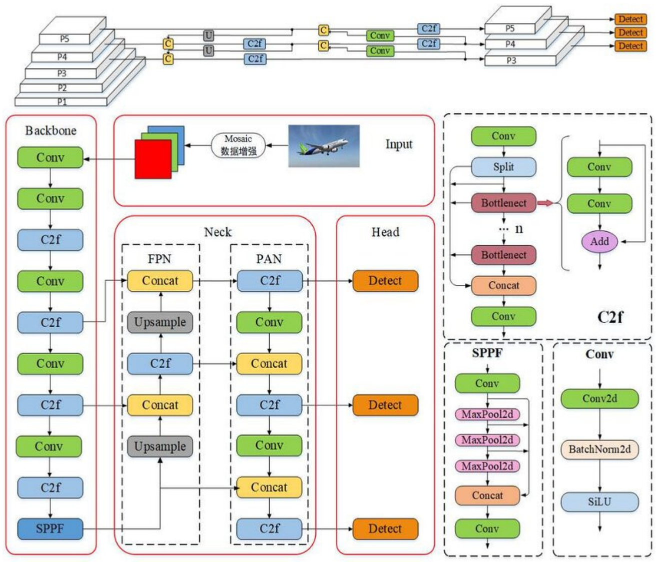

Figure 1. YOLOv8 Architecture

Backbone - Feature Extraction

The backbone of YOLOv8, an enhanced variant of CSPDarknet, is designed to extract multi-scale features from the input image with optimal computational efficiency. This architecture employs Conv Blocks as the fundamental units, comprising 2D Convolutional layers (typically 3x3 or 1x1) combined with Batch Normalization and the SiLU activation function, repeated to form feature layers (P1–P5). C2f, an improvement over YOLOv5's C3, integrates ResNet-like shortcuts and Bottleneck blocks to enhance representation capability without increasing model complexity. At the end of the backbone, SPPF (Spatial Pyramid Pooling Fast) utilizes MaxPooling with kernel sizes of 5x5, 9x9, and 13x13 to aggregate spatial features across multiple scales, thereby supporting multi-size object detection.

Neck - Multi-scale Feature Aggregation

The Neck of YOLOv8 functions as an intermediate stage, combining Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) to aggregate features from different layers of the Backbone. FPN conveys semantic information from deep layers to shallow layers, while PAN supplements spatial information from shallow layers to deeper ones, resulting in multi-scale feature maps. This mechanism enables the model to efficiently process objects of varying sizes, thereby enhancing its multi-scale detection capability.

Head - Prediction

The Head is responsible for the final prediction task, including bounding box localization, object classification, and confidence estimation, based on features derived from the Neck. This stage is critical in converting feature representations into visual outputs, suitable for real-time applications such as driver monitoring systems. A major advancement of the Head lies in its anchor-free architecture, which entirely replaces the anchor box mechanism used in previous versions. Instead of relying on manually tuned anchor boxes, each point on the feature map directly predicts the bounding box using a vector (l, t, r, b, conf, cls), where l, t, r, b represent the distances from the central point to the left, top, right, and bottom edges of the box, respectively. This approach simplifies the training pipeline, reduces dependency on hyperparameters, and improves generalization. Performance is further optimized with an 18% reduction in inference time, a 25% decrease in memory usage, and a 2.1% increase in mAP on the COCO dataset compared to YOLOv5. This not only accelerates processing but also enhances the efficiency of the Non-Maximum Suppression (NMS) procedure.

❖ ***Loss Function***

YOLOv8's loss function comprises three primary components: Box Loss, Classification Loss, and Distribution Focal Loss (DFL). The Box Loss, based on Complete Intersection over Union (CIoU), measures the accuracy of predicted bounding boxes with respect to ground-truth boxes. It accounts for overlap area, the distance between center points, and aspect ratio consistency. This is particularly effective for detecting small objects. The Classification Loss focuses on improving the accuracy of predicted object classes. Meanwhile, DFL is designed to address class imbalance in object detection tasks. Rather than predicting a fixed bounding box, DFL computes a probability distribution over the bounding box coordinates to account for spatial uncertainty. DFL enhances classification precision by concentrating probability mass around the precise locations of small objects such as eyes, improving detection reliability even under low-resolution conditions.

To provide a detailed mathematical representation, the overall loss function of YOLOv8 can be formulated as a weighted sum of its constituent components. The formulation begins with box loss, which measures the discrepancy between predicted and ground truth bounding box coordinates. Next, classification loss evaluates the deviation between predicted probabilities and actual labels using Binary Cross-Entropy (BCE). Finally, DFL is incorporated to optimize the confidence distribution, thereby improving the model's ability to distinguish between true and false detections. These components are integrated to form a comprehensive loss function that effectively drives the model's learning process.

Box loss in YOLOv8 utilizes Complete Intersection over Union (CIoU), which considers overlap area, the Euclidean distance between center points, and the aspect ratio between predicted and ground truth boxes. This can be mathematically defined as follows:

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho^2(\mathbf{b},\mathbf{b}^{gt})}{c^2} + \alpha v \quad (1)$$

Where IoU is Intersection over Union between predicted and ground-truth boxes, $\rho^2(\mathbf{b},\mathbf{b}^{gt})$ is Euclidean distance between the centers of the predicted box $\mathbf{b}$ and the ground-truth box $\mathbf{b}^{gt}$, c^2 is diagonal length of the smallest enclosing box, α is a Balancing factor between IoU and v, v: Measures aspect ratio difference.

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \quad (2)$$

Classification loss is based on Binary Cross-Entropy (BCE), a widely used loss function in binary classification problems, which quantifies the divergence between predicted probabilities and ground-truth labels. The formula is expressed as follows:

$$\mathcal{L}_{Classification} = -\frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} y_i \log(\hat{y}_i) \\ +(1-y_i)\log(1-\hat{y}_i)\end{bmatrix} \quad (3)$$

Where N is number of Predictions, y_i is ground truth label, $\hat{y}_i$ is predicted probability of class i.

Distribution Focal Loss (DFL) is designed to address data imbalance and enhance focus on difficult examples. The formula is defined as:

$$\mathcal{L}_{\mathrm{DFL}(\mathcal{S}_i,\mathcal{S}_{i+1})} = -\begin{pmatrix}(y_{i+1}-y)\log(\mathcal{S}_i) \\ +(y-y_i)\log(\mathcal{S}_{i+1})\end{pmatrix} \quad (4)$$

Where $\mathcal{S}_i, \mathcal{S}_{i+1}$ are predicted probabilities at grid points y_i and y_{i+1}, calculated from the softmax function, y is actual ground truth value to be predicted, y_i, y_{i+1} are two adjacent integer grid points surrounding y, $y_{i+1} - y$, $y - y_i$ are linear interpolation weights, with the closer point to y having a higher weight.

The overall loss function integrates these components with appropriate weighting factors to balance their influence:

$$\mathcal{L}_{total} = \lambda_1 \cdot \mathcal{L}_{CIoU} + \lambda_2 \cdot \mathcal{L}_{Classification} + \lambda_3 \cdot \mathcal{L}_{DFL} \quad (5)$$

Where $\lambda_1, \lambda_2, \lambda_3$ are hyperparameters that adjust the contribution of each component, typically tuned during training to optimize performance.

3.2. System design

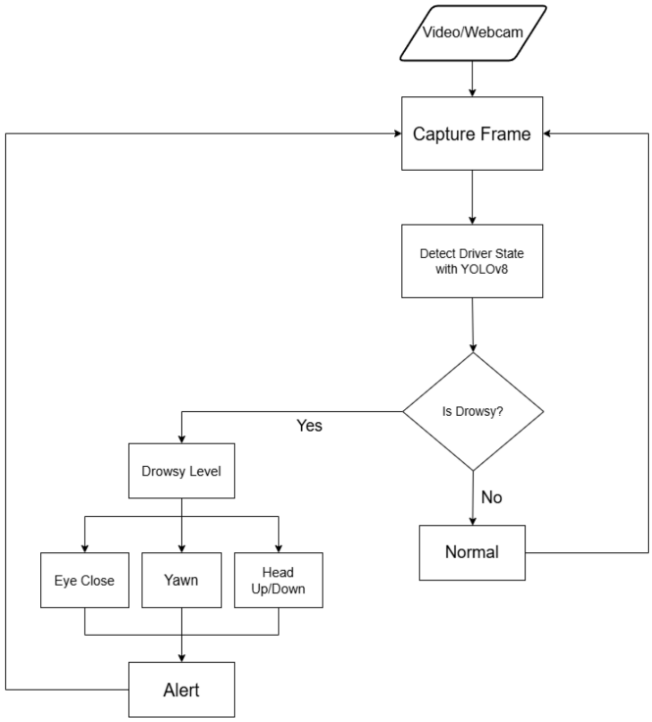

Figure 2. System design

The system is developed to monitor the driver's condition in real-time, thereby enabling timely alerts upon detecting signs of fatigue or drowsiness. The block diagram in Figure 2 is designed to illustrate the complete sequence of main processing steps, from system initialization to the issuance of alerts.

Specifically, the drowsiness detection system using YOLOv8 operates as follows:

The system initiates by acquiring visual input from an onboard video camera or webcam. It continuously captures individual frames for analysis. Each captured frame is processed by the YOLOv8 model to detect the driver's state.

The system evaluates whether the driver exhibits signs of drowsiness. If no such signs are detected, the driver's state is classified as "Normal", and the system resumes frame capturing for continued analysis.

If a drowsy state ("Drowsy") is detected, the system will analyze the drowsiness level ("Drowsy Level") based on three primary indicators:

• "Eye Close": assessing whether the driver's eyes are closed,

• "Yawn": detecting yawning behavior,

• "Head Up/Down": monitoring head movements indicative of sleepiness.

Upon analyzing these indicators, the system triggers an **alert** to awaken the driver, thereby mitigating the risk of accidents caused by drowsiness. The system then resumes frame capturing to maintain continuous monitoring.

4. RESULTS AND DISCUSSTION

4.1. Experimental dataset

To build and train the model, the Drowsiness Detection Xsriz dataset was utilized. The dataset consists of 6,391 images, which were divided into three subsets: train, validation, and test, with 9 labels including "eye_closed", "eye_closed_head_right", "eye_closed_head_left", "focused", "head down", "head up", "seeing_right", "seeing_left", "yawning". The data underwent preprocessing and data augmentation techniques to enhance diversity and ensure effective model training.

4.2. Experimental environment

The model training was conducted on the Kaggle platform, utilizing dual NVIDIA T4 GPUs (32 GB total VRAM, 16 GB per GPU) to train the YOLOv8 model. The system configuration included 4 CPU cores and 29 GB of RAM, enabling efficient data processing. The environment operated on Windows and employed the PyTorch and OpenCV libraries. The model was trained using the **YOLOv8s** architecture, with a batch size of 16 over the course of 100 epochs.

4.3. Experimental results

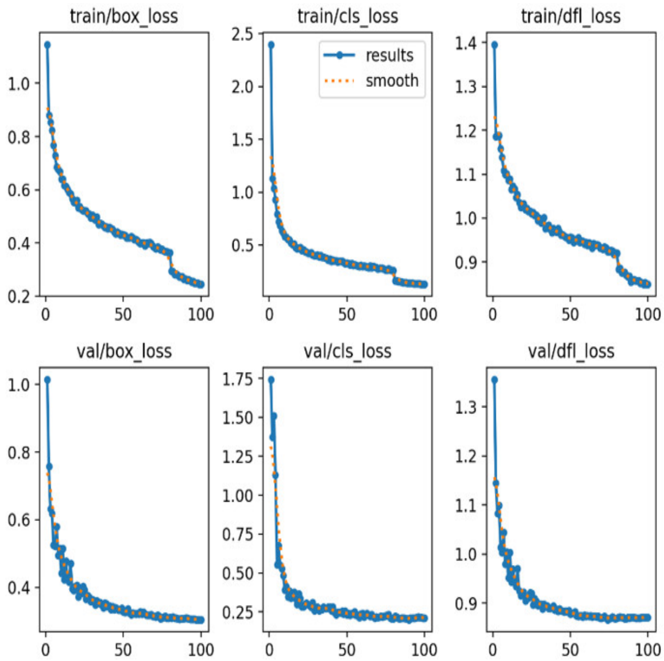

Figure 3. Loss Functions

Figure 3 presents graphs illustrating the evolution of loss functions during the training and validation of the YOLOv8 model, providing visual evidence of convergence when applied to driver drowsiness detection. The collection consists of six plots, including bounding box loss (box loss), classification loss (cls loss), and distribution focal loss (DFL loss) in both the training and validation phases. Each plot displays two lines: the blue line represents the actual recorded values, while the orange line denotes the smoothed values, tracked across 100 epochs.

Specifically, the bounding box loss curve shows a significant decrease from an initially high value to a substantially lower level during both training and validation, indicating the model's improved ability to localize bounding boxes around critical features such as the eyes, mouth, or entire face. This reduction reflects the model's capacity to learn more accurate alignment, minimizing overlap errors and adapting to low-resolution inputs - an essential factor for real-world detection of pupils and drowsiness cues. The classification loss curve also decreases over time, demonstrating enhanced discrimination between alert and drowsy states, such as distinguishing closed eyes from open eyes, although the rate of decline slows at times due to the subtlety of the cues involved. Similarly, the distribution focal loss exhibits a consistent downward trend, indicating the model's refined ability to optimize bounding box predictions, particularly in sensitive regions like the eyes and mouth.

At the beginning of training, high loss values are expected as the model begins learning features and spatial patterns associated with drowsiness. As training progresses, all curves drop sharply, and particularly near epoch 100, they begin to stabilize reflecting the law of diminishing returns, where additional training yields only marginal improvements. For optimal performance, loss functions should converge to lower values, ensuring that the model maintains a reliable balance among localization, classification, and drowsiness signal prediction. These results confirm YOLOv8's adaptability for real-time monitoring, even in diverse environmental conditions such as varying lighting and unstable camera angles.

Figure 4 presents the charts reflecting the performance metrics of the YOLOv8 model throughout the training process, providing a visual insight into its effectiveness in detecting driver drowsiness. The collection consists of four plots, including precision, recall, mAP at threshold 0.50 (mAP50), and mAP at threshold 0.50:0.95 (mAP50-95), recorded over 100 epochs. Each chart displays two curves: the blue line represents the actual results, and the orange line shows the smoothed values.

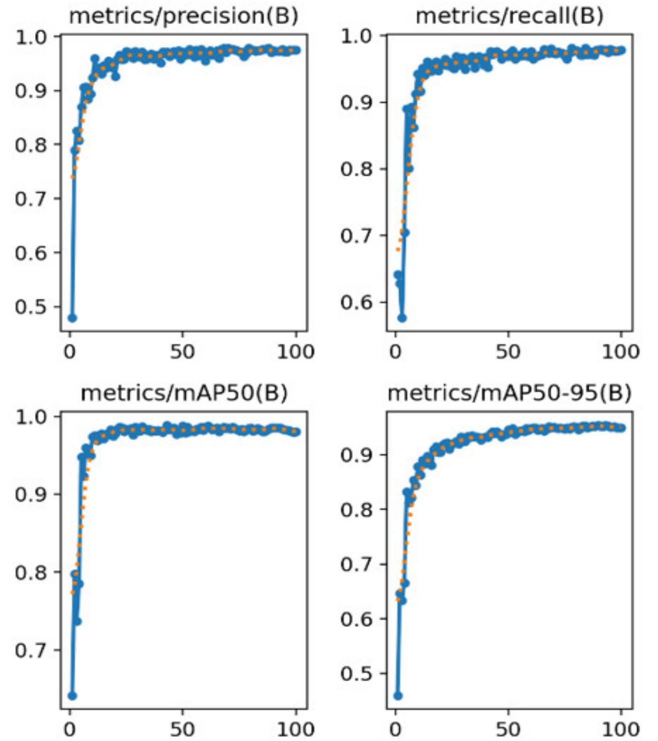

Figure 4. Model Performance Evaluation

Precision measures the proportion of correct drowsiness predictions out of the total predictions, indicating the reliability in recognizing fatigue-related signs; this curve rises from a low initial value to a near-high level, reflecting improvements in reducing false positives. Recall, or detection ability, assesses the proportion of actual drowsiness cases that were correctly identified compared to the total actual cases; this curve gradually increases, with slight fluctuations, suggesting the model gradually adapts to complex scenarios. mAP50 (mean Average Precision at IoU threshold 0.50) is the average precision metric over confidence thresholds, reflecting the alignment between predicted bounding boxes and ground truth at an IoU of 0.50; this curve steadily increases, indicating consistent performance at a moderate threshold. During training, mAP50 reached 98.6% at epoch 100, demonstrating high accuracy in matching predicted bounding boxes with ground truth labels at a medium confidence level. mAP50-95, the mean mAP across the IoU range from 0.50 to 0.95, evaluates performance over various confidence thresholds; this curve increases more slowly, suggesting that further improvement is needed at higher IoU thresholds.

In the initial epochs, the metrics started low as the model began to learn features related to drowsiness. Throughout the 100 epochs, the curves increased and maintained stable levels, demonstrating an effective learning process, particularly in identifying subtle features such as eye closure or yawning. This trend indicates YOLOv8's adaptability in handling real-world conditions such as uneven lighting or changing camera angles.

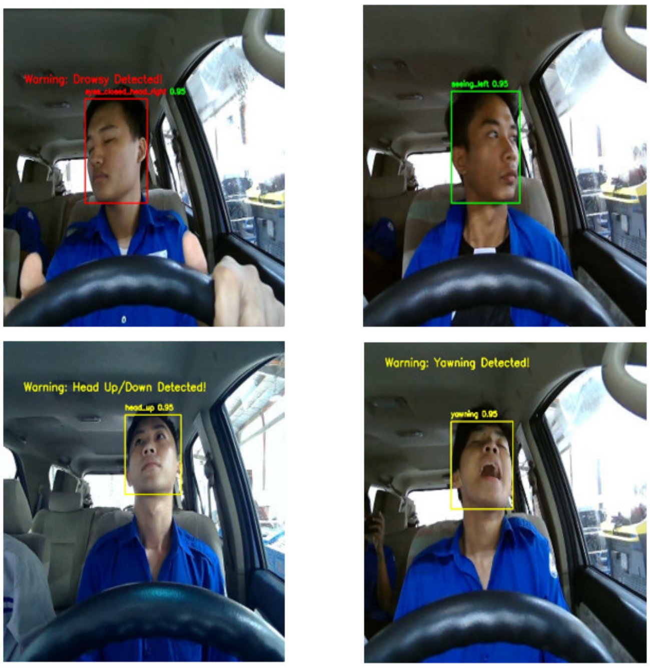

Figure 5. Result

After training, the model performs predictions of driver drowsiness states as presented in Figure 5. The YOLOv8 model identifies these indicators by analyzing facial states, such as the condition of the eyes, mouth, and nodding head.

5. CONCLUSIONS

The study proposes an effective solution for detecting driver drowsiness by employing the YOLOv8 model. Drowsiness state recognition is conducted by monitoring facial expressions of the driver extracted from video frames. The model achieves high Precision, Recall, and mAP50 in detecting fatigue indicators, confirming the effectiveness of the solution. The system was evaluated on the Drowsiness Detection Xsriz dataset and achieved an mAP50 of up to 98.6%, demonstrating outstanding recognition performance and the model's stability under experimental conditions.

Compared to the study by Krishna et al. using Vision Transformers and YOLOv5 [12], or the study by Balaji and colleagues employing YOLOv8 + CNN [11], the proposed model achieved a precision improvement of 1.2% higher than the YOLOv5 combined with ViT, and 13% higher than the previous YOLOv8-based study. In addition, this study monitors facial states based on nine labels, enabling clearer classification of facial expressions and improving recognition capability in complex real-world scenarios.

However, the system still presents several limitations. First, the current model mainly relies on facial data and may be affected by poor lighting conditions, drivers wearing masks, or unfavorable camera angles. Second, the system has not yet integrated other biometric indicators such as head movements, heart rate, or driving behavior to enhance reliability. Lastly, the system has only been tested on available datasets and has not been fully evaluated in real-world situations involving different environments and drivers.

In the future, development will focus on improving the stability and adaptability of the system under more complex real-world conditions. Potential solutions include integrating multimodal data sources such as IR sensors, infrared cameras, or biometric sensors; optimizing the model to reduce resource consumption on embedded devices; and conducting real-world testing in vehicles to comprehensively evaluate system effectiveness.

REFERENCES

[1]. Nguyen Thi Bich Nga, *The underlying cause behind severe and alarming traffic accidents.* 2024. Last accessed: 09/06/2025, URL: https://www.nguoiduatin.vn/nguyen-nhan-dang-sau-cac-vu-tai-nan-giao-thong-nghiem-trong-het-suc-bao-dong-204652574.htm.

[2]. Vu H. M., Tran T. H., Dang A. K., Hoang T. N., Nguyen C. T., Nguyen H. L. T., Latkin C. A., Ho C. S. H., Ho R. C. M., "Sleep disorders among patients suffering from road traffic injuries in an urban setting of Vietnam: An exploratory study," *Scientific Reports*, 13(1), 2023. https://doi.org/10.1038/s41598-023-38693-7.

[3]. National Safety Council, 2024, Last accessed: 09/06/2025, URL: https://www.nsc.org/

[4]. Girshick R., "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision* (ICCV), Santiago, Chile, December 2015. https://doi.org/10.1109/ICCV.2015.169.

[5]. Joseph R., Santosh D., Ross G., Ali F., "You only look once: Unified, Real-Time Object Detection," *arXiv (Cornell University),* 2015. https://doi.org/10.48550/arxiv.1506.02640

[6]. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., Berg A. C., "SSD: Single Shot MultiBox Detector," in *Lecture notes in computer science*, 21-37, 2016. https://doi.org/10.1007/978-3-319-46448-0_2

[7]. Yaseen M., "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," *arXiv (Cornell University),* 2024. https://doi.org/10.48550/arxiv.2408.15857

[8]. Aravind A., Agarwal A., Jaiswal A., Panjiyara A., Pm M. S., "Fatigue detection system based on eye blinks of drivers," *International Journal of Engineering and Advanced Technology,* 8(5s), 72–74, 2019. https://doi.org/10.35940/ijeat.e1015.0585s19

[9]. Phan A., Trieu T., Phan T., "Driver drowsiness detection and smart alerting using deep learning and IoT," *Internet of Things*, 22, 100705, 2023. https://doi.org/10.1016/j.iot.2023.100705

[10]. Ahmed M. I. B., Alabdulkarem H., Alomair F., Aldossary D., Alahmari M., Alhumaidan M., Alrassan S., Rahman A., Youldash M., Zaman G., "A Deep-Learning approach to driver drowsiness Detection," *Safety,* 9(3), 65, 2023. https://doi.org/10.3390/safety9030065

[11]. Balaji C. G., Krishnaa V. R. S., Sundar S. S., Kumar S. R., "Driver Safety Advancements: Drowsiness Detection with YOLO V8," in *Communications in Computer and Information Science*, 162-175, 2025. https://doi.org/10.1007/978-3-031-82383-1_13

[12]. Krishna G. S., Supriya K., Vardhan J., K M. R., "Vision Transformers and YOLOV5 based Driver Drowsiness Detection Framework," *arXiv (Cornell University),* 2022. https://doi.org/10.48550/arxiv.2209.01401

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

[14]. Albadawi Y., AlRedhaei A., Takruri M., "Real-Time Machine Learning-Based driver drowsiness detection using visual features," *Journal of Imaging*, 9(5), 91, 2023. https://doi.org/10.3390/jimaging9050091