DRIVER STATE MONITORING AND ASSISTED INTERVENTION BASED ON DEEP LEARNING

GIÁM SÁT TÌNH TRẠNG NGƯỜI LÁI VÀ HỖ TRỢ CAN THIỆP CHỦ ĐỘNG DỰA TRÊN CÔNG NGHỆ HỌC SÂU

Nguyen Le Chau Thanh*, Phung Minh Tung, Kieu Quoc Long, Ho Thang Lanh, Trinh Phan Minh Vu

The University of Danang – University of Technology and Education, Vietnam

*Corresponding author: nlcthanh@ute.udn.vn

(Received: March 04, 2026; Revised: May 12, 2026; Accepted: May 20, 2026)

DOI: 10.31130/ud-jst.2026.24(5A).123E

Abstract - Driver drowsiness is one of the common factors to traffic accidents, particularly on long-distance routes and highways. In recent years, numerous studies on driver state monitoring based on deep learning have been proposed, mainly focusing on improving recognition accuracy and adaptability to various environmental conditions. However, most of these studies remain at the warning level, while recognition to support proactive risk mitigation and driver's safety-oriented support remains limited. This paper proposes a driver state monitoring system using deep learning models to monitor eye states and yawning behaviors of the driver. Based on the recognition outcomes, a lightweight intervention support mechanism is implemented to reduce potential risks for drivers.

Key words - Deep Learning; Driver Monitoring; Assited Intervension; Convolution Neural Network

Tóm tắt - Tình trạng buồn ngủ của tài xế là những nguyên nhân phổ biến gây ra các vụ tai nạn giao thông, đặc biệt là trên các tuyến đường dài và cao tốc. Trong những năm gần đây, nhiều nghiên cứu giám sát tình trạng người lái dựa trên công nghệ học sâu được đề xuất, chủ yếu tập trung vào việc cải thiện độ chính xác nhận diện và khả năng thích ứng nhiều điều kiện môi trường. Tuy nhiên, phần lớn các nghiên cứu này vẫn dừng lại ở mức cảnh báo, trong khi việc nhận diện để hỗ trợ giảm thiểu rủi ro chủ động và an toàn cho người lái vẫn còn hạn chế. Bài báo này đề xuất một hệ thống giám sát trạng thái người lái sử dụng mô hình học sâu để giám sát trạng thái mắt và hành vi ngáp của tài xế. Dựa vào kết quả nhận diện, một cơ chế hỗ trợ can thiệp mức độ nhẹ được triển khai nhằm giảm thiểu rủi ro cho người lái.

Từ khóa – Học sâu; Giám sát người lái; Hỗ trợ can thiệp; Mạng nơ-ron tích chập

1. Introduction

Driver drowsiness is often underestimated, although it is considered one of the leading causes of traffic accidents resulting in injuries and fatalities worldwide. According to the AAA Foundation report, drowsy driving accounts for approximately 8.8–9.5% of all investigated crashes, with those leading to significant human and property damage or injuries representing 10.6–10.8% [1]. In the United States, the National Highway Traffic Safety Administration (NHTSA) recorded approximately 684 fatalities related to drowsy driving in 2021 [2].

Drowsy driving occurs silently; a loss of attention for only a few seconds may lead to severe consequences. Contributing factors such as night-shift work [3], sleep disorders, and prolonged driving duration [4] have been studied and shown that accumulated fatigue significantly reduces the ability to maintain alertness, thereby increasing the risk of traffic accidents.

In recent years, the demand for driver state monitoring systems has continuously increased. In general, existing monitoring approaches can be categorized into two main groups: invasive and non-invasive methods. Although invasive methods can directly assess the physiological state of the driver, they cause inconvenience to users [5]. In contrast, non-invasive methods based on computer vision are considered more suitable for practical applications due to minimal interference with the driver. However, traditional methods relying on handcrafted features still encounter significant limitations under complex environmental conditions [6].

The advancement of deep learning technology, particularly Convolutional Neural Networks (CNNs), has opened a more effective approach due to their capability of automatic feature extraction in tasks such as facial state recognition and expression classification [7]. Nevertheless, most existing studies mainly focus on improving recognition accuracy and remain at the warning level, while leveraging recognition results to provide proactive and safety-oriented assistance remains limited.

Motivated by these issues, this study proposes a CNN-based driver monitoring system through eye state classification and yawning behavior detection to evaluate driver status. Based on the assessment results, an assisted intervention module is integrated under assumed scenarios to reduce risks to life and property.

2. Overall System Architecture

This section presents the overall architecture of the proposed driver drowsiness monitoring and assisted intervention system. The system focuses on utilizing computer vision combined with deep learning to monitor driver state while providing graded intervention mechanisms to minimize risk.

Overall, the system consists of four main functional blocks: (i) Image acquisition from an in-cabin RGB camera; (ii) a deep learning-based driver state

classification module; (iii) a driver state evaluation module; (iv) a graded assisted intervention module. This architecture clearly separates the recognition stage, state inference stage, and decision-making stage.

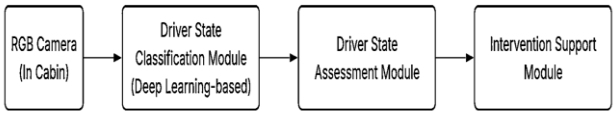

Figure 1. *Overview of the proposed system*

Specifically, the deep learning-based monitoring module employs lightweight CNN models to classify eye states and detect yawning behavior from Regions of Interest (ROI) extracted from the face. Recognition results are then forwarded to the driver state evaluation module. The driver state is categorized into multiple levels based on accumulated information within sliding time windows, allowing temporal assessment of driver alertness.

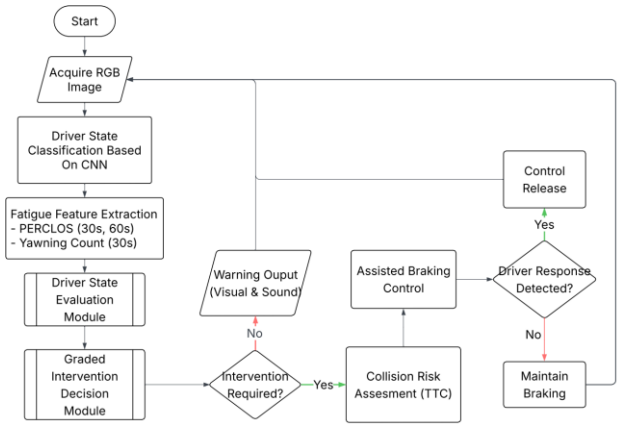

Figure 2. *Flowchart of the driver monitoring and assisted intervention system*

After determining the drowsiness levels, the system activates intervention strategies corresponding to each level. Support measures are sequentially triggered, starting with mild warnings such as audio and visual signals and escalating to higher intervention levels when the driver is detected to be incapable of maintaining vehicle control under assumed scenarios. This graded approach aims to prevent unnecessary inconvenience while keeping the driver at the center of intervention decisions.

2.1. Driver State Classification Module

2.1.1. Facial ROI Extraction

The input data are collected using an RGB camera installed inside the vehicle cabin. This approach ensures non-invasiveness and does not interfere with the driver.

To achieve accurate eye and mouth localization, the proposed system employs MediaPipe FaceMesh with 468 3D facial landmarks [8]. Unlike traditional Haar cascade detectors [9] or Dlib-based facial landmark approaches [10], which may degrade under varying illumination and large head poses, FaceMesh estimates a dense 3D facial mesh that enables more stable ROI extraction and improved localization accuracy [8].

However, experimental observations show that when the camera is positioned at a distance of 50–60 cm from the driver's face, the ROI generated directly from FaceMesh landmarks does not align well with the image regions used during model training. This domain shift leads to reduced prediction accuracy. Therefore, this study proposes an ROI adjustment strategy based on expanding the bounding box [11] extracted from facial landmarks.

Figure 3. *Distance between the camera and the driver*

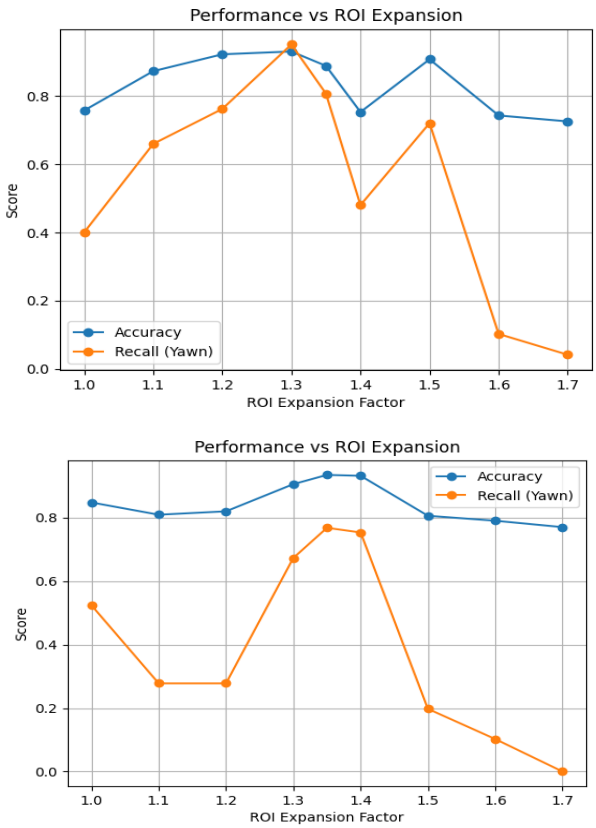

Figure 4. *Effect off ROI expansion factor on facial state prediction performance*

The initial facial bounding box $L = \{p_i = (x_i, y_i)\}_{i=1}^{n}$ is determined from landmark limits follows: $x_{min} = \min(x_i)$, $x_{max} = \max(x_i)$, $y_{min} = \min(y_i)$, $y_{max} = \max(y_i)$. The width and height of the bounding box are computed as: $w = x_{max} - x_{min}$, $h = y_{max} - y_{min}$

To incorporate additional image context and reduce the domain shift between training and real-world data, the

bounding box is expanded along both dimensions using scaling factors α và β.

The expansion is defined as:

$$\Delta x = \alpha w, \quad \Delta y = \beta h$$

The coordinates of the expanded bounding box are computed as:

$$x'_{min} = \max(0, x_{min} - \Delta x), \; x'_{max} = \min(W, x_{max} + \Delta x)$$

$$y'_{min} = \max(0, y_{min} - \Delta y), \; y'_{max} = \min(H, y_{max} + \Delta y)$$

Where H and W denote the height and width of the input image, respectively.

The expansion parameter was evaluated using two independent videos to determine its impact on precision and recall. Both datasets show consistent performance trends. The highest recall is observed when the ROI expansion factor lies between 1.3 and 1.4. Precision also reaches its maximum within this range for both datasets, indicating an optimal trade-off between precision and recall.

When the expansion factor exceeds 1.5, recall decreases due to domain shift. Based on these results, the expansion factor $\alpha = \beta = 1.35$ is selected for the proposed system to ensure stable performance across experimental videos.

2.1.2. Eye Region Extraction Based on Facial Landmarks

After determining the facial ROI, the system extracts the eye regions for eye state classification. The eye region is a critical parameter in evaluating driver alertness, particularly based on blink frequency and prolonged eye closure duration.

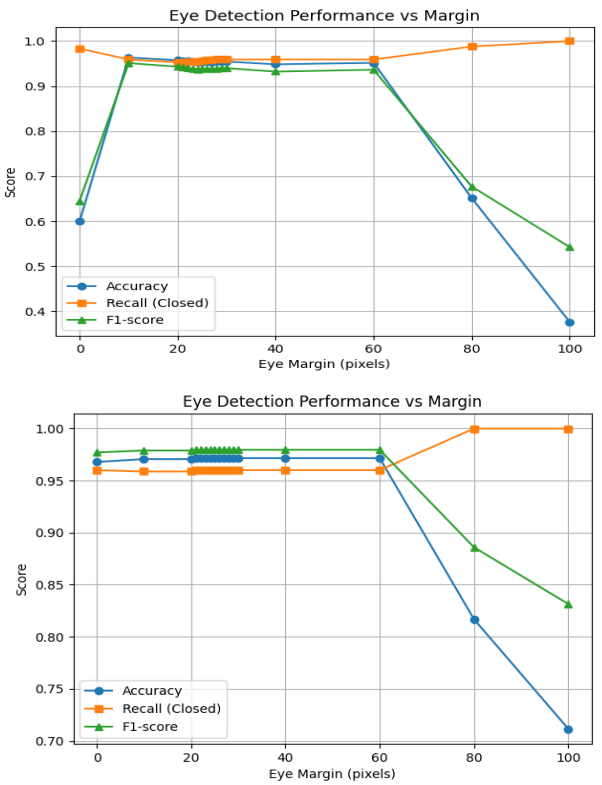

Figure 5. *Effect of ROI expansion factor on eye state prediction performance*

In this study, the eye regions are extracted from MediaPipe FaceMesh landmarks. Each eye is defined by eight landmark points, enabling precise localization of the upper eyelid, lower eyelid, and eye corners in each frame. Similar to the facial ROI, the eye ROIs are expanded using a margin parameter applied to the bounding box in four directions.

The expanded bounding box is computed as:

$$x'_{min} = \max(0, x_{min} - margin), x'_{max} = \min(W, x_{max} + margin)$$

$$y'_{min} = \max(0, y_{min} - margin), y'_{max} = \min(H, y_{max} + margin)$$

Where:

x_{min}, x_{max}, y_{min}, y_{max}: denote the initial eye bounding box coordinates determined from the landmarks;

W, H: denote the frame width and height, respectively;

margin: the expansion parameter applied in four directions.

Experimental evaluation with different margin values was conducted to assess impacts on precision, recall, and F1-score. When the margin is too small (<20 px), the extracted ROI becomes narrow and noisy. When the margin is too large (>60 px), irrelevant background regions are included, causing domain shift and performance degradation. A margin range of 22-26 px maintains high recall while precision and F1-score reach stable saturation. Therefore, margin = 24 px is selected.

2.1.3. Proposed CNN Architecture and Training Methodology

After ROI extraction and adjustment, eye-state and yawning behavior recognition are performed using two independent lightweight CNN models. Although both tasks are related to driver fatigue analysis, previous studies have reported that shared multi-task learning frameworks may suffer from task interference and feature competition, leading to degraded task-specific representation capability [12-14]. Therefore, two independent CNN models with identical architectures are adopted to improve classification robustness and training stability.

To ensure computational efficiency suitable for real-time in-cabin deployment, a lightweight CNN architecture is proposed. The network processes RGB input images resized to 255×255 pixels. The main feature extraction consists of five sequential convolutional blocks, where each block applies a 3×3 convolutional layer with a ReLU activation function followed by a 2×2 max-pooling layer for spatial downsampling. The number of filters progressively increases from 32 to 128. Following feature extraction, the feature maps are flattened and connected to a fully connected layer with 128 neurons. To improve generalization and reduce overfitting, Batch Normalization and a Dropout layer with a rate of 0.3 are employed before the final Softmax classification layer. The proposed architecture contains fewer than one million trainable parameters, providing a balance between classification accuracy and computational efficiency for embedded driver-monitoring applications.

The models were trained using a dataset comprising approximately 3,000 facial images collected from publicly available datasets and refined through ROI extraction and preprocessing. The dataset includes driver facial states associated with eye closure and yawning behavior and was partitioned into 80% training, 10% validation, and 10% testing subsets. To improve robustness despite the limited dataset size, on-the-fly data augmentation was applied exclusively to the training set, including random horizontal/vertical flipping and rotation operations.

During training, the models were optimized using the Adam optimizer and Sparse Categorical Crossentropy loss function with a batch size of 32. An Early Stopping mechanism with a patience of 15 epochs was employed to monitor validation loss and restore the best-performing weights. Experimental results demonstrated stable convergence behavior without significant overfitting between training and validation stages.

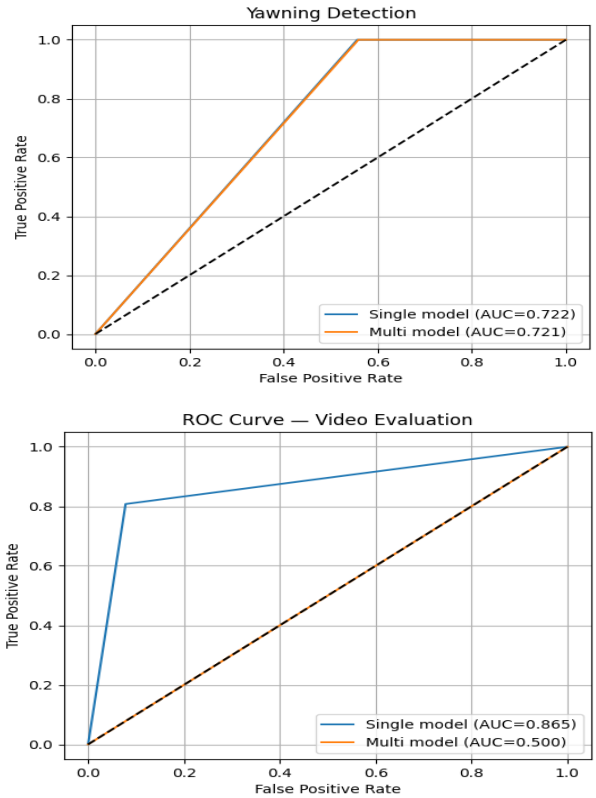

Figure 6. *Performace comparision between independent model and multi-task model*

The ROC analysis shown in Figure 6 further supports the effectiveness of the proposed single-task learning strategy. For yawning detection, the independently trained CNN achieved an AUC score of 0.722, slightly outperforming the multi-task model. More importantly, during video evaluation under continuous monitoring conditions, the single-task model achieved a significantly higher AUC score of 0.865, while the multi-task model exhibited unstable performance with an AUC value close to random classification. These results indicate that separating eye-state and yawning recognition into two dedicated CNN models improves temporal robustness and classification reliability in practical driver-monitoring scenarios.

The eye-state model outputs class open and closed eyes, while the yawning model classifies yawning and non-yawning behavior. Individual frame-level predictions are subsequently aggregated through sliding-window analysis to reduce temporary misclassification and improve temporal robustness during driver-state evaluation.

2.2. Driver State Evaluation Module

After classification by CNN, driver state is evaluated using the PERCLOS (Percentage of Eye Closure) index, defined as the proportion of time during which the driver's eyes are at least 80% closed [16]. In the proposed system, PERCLOS is computed over two sliding windows with durations T = 30 s and T = 60 s. The resulting value are subsequently utilized to determine the corresponding intervention levels.

$$\text{PERCLOS} = \frac{N_{Closed}}{N}$$

Where, N_{Closed}: the number of frames classified as closed eyes; N: the total number of frames within duration T.

Based on the synthesis of previous studies [17-21], when the PERCLOS value exceeds the selected threshold for state classification, the driver is considered to exhibit signs of drowsiness or reduced alertness. The aforementioned studies have employed PERCLOS thresholds in the range of 0.12-0.30 to identify significant drowsiness. Accordingly, the proposed system adopts a PERCLOS threshold of 0.10 for a sliding window of T = 30s to enable early detection of alertness degradation. In addition, a threshold of PERCLOS = 0.30 is applied for a sliding window of T = 60s to identify pronounced drowsiness or severely reduced driving capability. This time-based approach allows the system to detect early signs of fatigue, improve reliability, and reduce the influence of random physiological blinking.

In addition to PERCLOS-based evaluation, yawning behavior is incorporated as a complementary indicator to enhance early detection of driver fatigue. Behavioral cues such as yawning and eye closure have been widely used in driver monitoring studies as observable signs of reduced alertness [15].

Since yawning behavior is temporally sparse rather than continuously observable, its frequency is evaluated within the same sliding time window and defined as:

$$f_{yawn} = \frac{N_{yawn}}{T}$$

Where, N_{yawn} is the number of classified yawning events within the duration T.

Instead of using a rigid threshold, the proposed system utilizes yawning frequency as a supplementary factor to increase the sensitivity of early fatigue detection. Specifically, the presence of yawning events within a short duration may reinforce the activation of early warning, even when the PERCLOS value has not yet exceeded the predefined threshold. This mechanism enables the system

to identify early-stage fatigue prior to significant PERCLOS elevation.

However, due to its intermittent and less temporally stable nature, yawning behavior is not utilized as the primary trigger for safety-critical intervention. Therefore, PERCLOS remains the principal metric for intervention decisions, ensuring robustness and minimizing false-positive activations.

The proposed intervention framework is modular and primarily depends on fatigue-related behavioral indicators rather than a specific recognition architecture, allowing integration with alternative driver-monitoring approaches capable of providing equivalent fatigue-state estimations.

2.3. Graded Assisted Intervention Module

After the driver's state is evaluated by the state assessment module, the graded assisted intervention module is activated to mitigate accident risks associated with reduced driver alertness. The proposed framework prioritizes early warning before active intervention.

The intervention strategy primarily relies on PERCLOS values computed within sliding time windows, while yawning frequency is incorporated as a complementary factor to enhance early-stage detection. The system defines two graded levels:

- Early warning level: When the PERCLOS value exceeds 0.10 within a sliding window of T=30s, the driver is considered to exhibit reduced alertness. In addition, frequent yawning events within the same temporal window may reinforce the activation of visual and auditory warnings even when the PERCLOS value remains close to the predefined threshold. This level aims to help the driver regain alertness through early non-invasive notifications.

- Intervention support level: When the PERCLOS value exceeds 0.30 within a sliding window of T=60s, the system identifies a high risk of prolonged drowsiness or loss of vehicle control capability. At this stage, the intervention module performs a final verification through intensified warnings combined with early-stage alerts. If no driver response is detected, the assisted intervention mechanism is activated to reduce potential risks.

3. Proposed Assisted Intervention Method

When the assisted intervention module is activated, the proposed system employs an intervention strategy based on the Time-To-Collision (TTC) index to estimate the level of danger in the forward traffic situation and make decisions to bring the vehicle to a safe stop [22, 23].

$$t_{TTC} = \frac{d_{rel}}{v_{rel}},\ v_{rel} = v_{ego} - v_{lead}$$

Where,

d_{rel}: the relative distance between the ego vehicle and the preceding vehicle;

v_{rel}: the relative velocity;

v_{ego}: the velocity of the ego vehicle;

v_{lead}: the velocity of the preceding vehicle.

The vehicle stopping time is computed as:

$$t_{stop} = \frac{v_{rel}}{|j|}$$

Where j denotes the braking acceleration of the vehicle

Based on t_{TTC} value, the system performs staged intervention through braking phases PB_1, PB_2, and FB with progressively increasing deceleration levels across each phase, ranging from Partial Braking (PB) to Full Braking (FB), according to the control logic $t_{TTC} < t_{stop}$ & $t_{TTC} < 0$ in order to reduce vehicle speed and ensure that the vehicle stops before a collision occurs.

The system is evaluated using assessment scenarios selected according to the Euro NCAP recommendations [24] and is implemented and validated for effectiveness within the MATLAB/Simulink simulation environment.

3.1. Simulation Scenarios

The simulation scenarios used to evaluate the effectiveness of the proposed assisted intervention strategy in ensuring driver safety are selected based on the Euro NCAP recommendations [24], including CCRs (Car-to-Car Rear Stationary), CCRm (Car-to-Car Rear Moving), and CCRb (Car-to-Car Rear Braking) scenarios.

It should be noted that the simulations are conducted under simplified conditions to focus on evaluating the effectiveness of the proposed intervention strategy. Therefore, environmental and weather effects, sensor uncertainties, and preceding-vehicle trajectory instability are neglected, while a 100% overlap between vehicles is assumed.

3.2. Validation of the effectiveness of the proposed method in the MATLAB/Simulink Simulation Environment

The vehicle selected for simulation is characterized by dynamic parameters chosen based on representative values of passenger sedan vehicles. This representative parameter selection is commonly adopted in studies on vehicle safety systems [24].

Table 1. *Model Parameters*

Component	Parameter	Value	Unit
Vehicle mass	m	1575	kg
Wheelbase	L	2800	mm
Distance from center of gravity to front axle	a	1450	mm
Distance from center of gravity to rear axle	b	1350	mm
Braking acceleration PB_1	j_{PB1}	3.8	m/s^2
Braking acceleration PB_2	j_{PB2}	5.3	m/s^2
Braking acceleration FB	j_{FB}	8.0	m/s^2

3.2.1. Dynamic Model of the Simulated Vehicle

The rotational motion of the vehicle in three-dimensional space is described by three Euler angles: roll (ϕ), pitch (θ) and yaw (ψ), corresponding to rotations about the longitudinal, lateral, and vertical axes attached to the vehicle body [25, 26].

The roll angle describes the lateral tilting motion of the vehicle body about the longitudinal axis and is commonly associated with rollover risk during cornering maneuvers [25]. However, under the assumed straight-line braking scenario on a flat road surface, the lateral acceleration is negligible and therefore is disregarded.

The pitch angle describes the rotational motion of the vehicle body about the lateral axis. During braking, longitudinal load transfer from the rear axle to the front axle may induce a nose-down pitching motion [26]. However, since the proposed study focuses on straight-line braking scenarios, pitch dynamics are assumed negligible because they have limited influence on the longitudinal vehicle trajectory and braking performance evaluation.

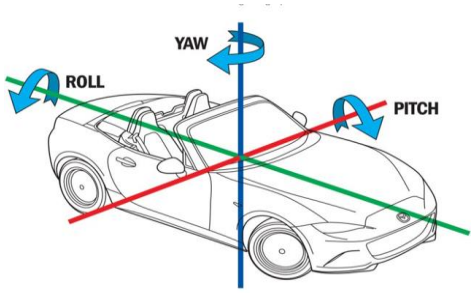

Figure 8. *Dynamic model of the vehicle*

In contrast, yaw motion directly reflects changes in the vehicle heading within the road plane [26]. Directional instability phenomena such as oversteer, understeer, or vehicle spin are manifested through variations in the yaw moment and yaw rate [27].

Accordingly, this study adopts a planar two-degree-of-freedom vehicle dynamic model that considers only lateral motion and yaw motion, while neglecting roll and pitch degrees of freedom, in order to evaluate the directional stability of the vehicle during intervention, especially under high deceleration braking phases.

3.2.2. Car-to-Car Rear Stationary (CCRs) Scenario

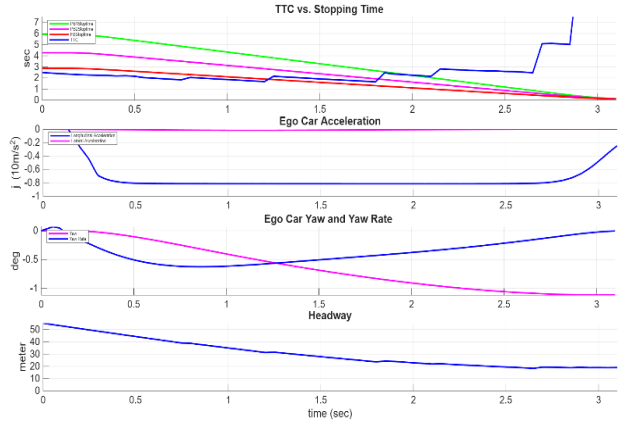

Figure 9. *Vehicle kinematic responses in the CCRs scenario*

In this scenario, the initial distance between the ego vehicle and the target vehicle is set to 60 m, and the initial velocity of the ego vehicle is 22.22 m/s (80 km/h). The target vehicle is assumed to be stationary. At the beginning of the simulation, the $t_{TTC} \approx 2.7$s, which is lower than the PB2 activation threshold of 4.2 s, indicating a medium-to-high risk situation when the driver loses control and the safe distance is not maintained.

The simulation results in the CCRs scenario indicate that the proposed system is capable of timely detection and intervention when the driver loses control. The longitudinal braking acceleration reaches -8 m/s², while the lateral dynamic indicators remain at low levels, and the vehicle maintains directional stability. The vehicle is able to decelerate and come to a stop before reaching the obstacle, with no collision occurring.

3.2.3. Car-to-Car Rear Moving (CCRm) Scenario

In the CCRm scenario, the ego vehicle travels at 22.22 m/s (80 km/h), while the target vehicle moves at a lower speed of 16.67 m/s (60 km/h), with an initial separation distance of 60m. The simulation results demonstrate that the proposed system enables the vehicle to stop safely without collision while maintaining directional stability during braking. These results confirm that the system satisfies the safety requirements under the tested scenario.

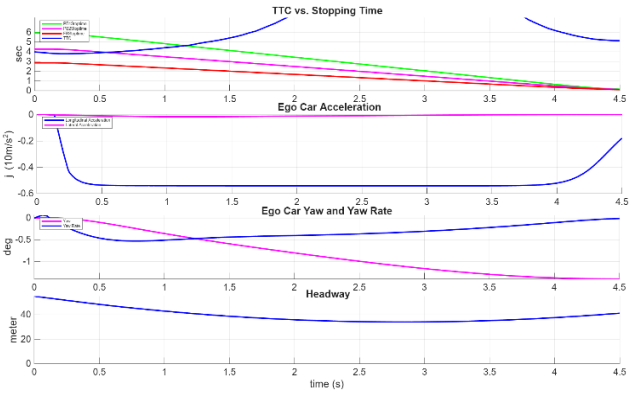

Figure 10. *Vehicle kinematic responses in the CCRm scenario*

3.2.4. Car-to-Car Rear Braking (CCRb) Scenario

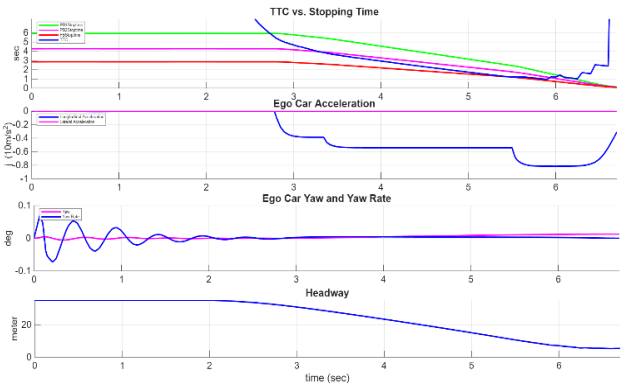

Figure 11. *Vehicle kinematic responses in the CCRb scenario*

In the CCRb scenario, the initial velocities of both vehicles are 13.89 m/s (50 km/h), the initial distance is 40 m, and the braking acceleration of the target vehicle is −6 m/s². The proposed system enables the ego vehicle to stop and avoid collision with the preceding vehicle during the simulation. The results show that the system meets the safety requirements, as no collision occurs and the vehicle remains stable during braking.

3.2.5. Braking Forces Generated at the Brake Mechanisms

The proposed system is evaluated under assumed scenarios with severe conditions requiring high braking deceleration. The braking force distributed to the front and rear axles is calculated as follows [28, 29]:

$$F'_F = \frac{Gb}{2L}(1+\frac{j_{max}h_g}{gb})\varphi \ , \ F'_R = \frac{Ga}{2L}(1-\frac{j_{max}h_g}{ga})\varphi$$

Where, G = mg is the vehicle weight;

a, b: the distances from the center of gravity to the front and rear axles, respectively;

h_g: the height of the center of gravity;

L: the wheelbase;

j_{max}: the braking deceleration;

g: the gravitational accleration;

φ: the tire-road adhension coefficient.

The simulation results show that when the vehicle transitions to maximum braking, the braking force distributed to the front axle increases significantly. This trend is consistent with the characteristic longitudinal load transfer toward the front axle during hard braking, thereby reflecting the rationality of the brake force distribution under high deceleration conditions.

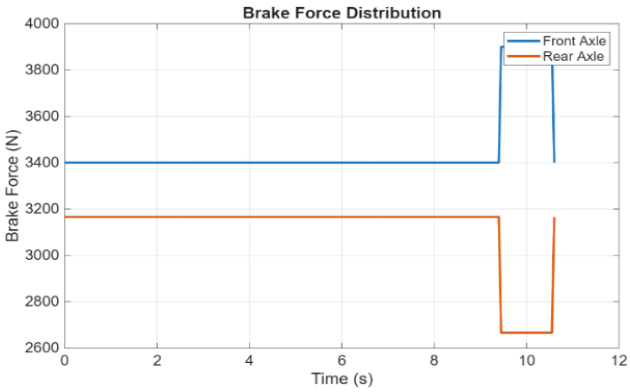

Figure 12. *Force Distribution at the Brake Mechanisms*

4. Conclusion

This paper has presented and developed a driver state monitoring system integrated with a proactive intervention capability based on deep learning technology. The main results achieved in this study can be summarized as follows. First, the study identified an optimal facial ROI expansion factor in the range of 1.3–1.4, along with an eye margin of 24 px, enabling the CNN model to maintain classification stability and improve accuracy. Second, the system performs drowsiness classification based on accumulated PERCLOS values over 30 s and 60 s sliding windows, allowing early detection of alertness degradation before activating the intervention mechanism. Third, experimental scenarios conducted in accordance with the Euro NCAP standards demonstrate that the TTC-based intervention strategy enables the vehicle to stop safely while maintaining directional stability under emergency braking conditions.

Although promising results have been obtained, several limitations remain to be addressed. Specifically, the size of the training dataset is still limited; therefore, further dataset expansion, data augmentation techniques, and the integration of IR cameras are necessary to improve model prediction performance under various lighting and weather conditions. In addition, the current intervention mechanism focuses solely on longitudinal control and does not consider lateral dynamics. Hence, the incorporation of trajectory planning or evasive control algorithms is required to enhance collision mitigation capability in situations where braking alone is insufficient.

Overall, the simulation results confirm the feasibility of the proposed intervention approach in mitigating risks associated with reduced driver alertness.

Future work will focus on implementing the system on real-time embedded platforms and conducting experiments on real vehicles under complex traffic conditions, aiming toward practical deployment of the proposed solution.

REFERENCE

[1] AAA Foundation for Traffic Safety, “Prevalence of drowsy driving crashes: Estimates from a large-scale naturalistic driving study,” *AAA Foundation for Traffic Safety*, 2018. [Online]. Available: https://aaafoundation.org/prevalence-drowsy-driving-crashes-estimates-large-scale-naturalistic-driving-study/ [Accessed Jan. 1, 2026].

[2] National Highway Traffic Safety Administration, “Countermeasures that work: Drowsy driving,” *National Highway Traffic Safety Administration,* 2020. [Online]. Available: https://www.nhtsa.gov/book/countermeasures-that-work/drowsy-driving [Accessed Jan. 1, 2026].

[3] L. Di Milia, “Shift work, sleepiness and long distance driving,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 9, no. 4, pp. 278–285, 2006. https://doi.org/10.1016/j.trf.2006.01.006

[4] M. L. Lee *et al*., “High risk of near-crash driving events following night-shift work,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 1, pp. 176–181, 2016. https://doi.org/10.1073/pnas.1510383112

[5] A. Sahayadhas, K. Sundaraj, and M. Murugappan, “Detecting driver drowsiness based on sensors: A review,” *Sensors*, vol. 12, no. 12, pp. 16937–16953, 2012. https://doi.org/10.3390/s121216937

[6] I. Nasri, M. Karrouchi, K. Kassmi, and A. Messaoudi, “A review of driver drowsiness detection systems: Techniques, advantages and limitations,” *arXiv preprint*, arXiv:2206.07489, 2022. https://doi.org/10.48550/arXiv.2206.07489

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25 (NeurIPS)*, Lake Tahoe, NV, USA, 2012, pp. 1097–1105.

[8] Y. Kartynnik, A. Ablavatski, I. Grishchenko et al., “Real-time facial surface geometry from monocular video on mobile GPUs,” *arXiv preprint*, arXiv:1907.06724, 2019. https://doi.org/10.48550/arXiv.1907.06724

[9] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proc. IEEE Conf. Computer Vision and Pattersn Recognition (CVPR)*, Kauai, HI, USA, 2001, pp. 511–518.

[10] V. Kazemi and J. Sullivan, “One millisecond face alignment with an ensemble of regression trees,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, 2014, pp. 1867–1874. https://doi.org/10.1109/CVPR.2014.241

[11] C. Nogueira, L. Fernandes, J. N. D. Fernandes, and J. S. Cardoso, “Explaining bounding boxes in deep object detectors using post hoc

methods for autonomous driving systems," *Sensors*, vol. 24, no. 2, article 516, 2024. https://doi.org/10.3390/s24020516

[12] M. M. Abdelsamie, S. S. Azab, and H. A. Hefny, "Deep multi-task learning: A review of concepts, methods, and cross-domain applications," *International Journal of Data Science and Analytics*, vol. 21, article 77, 2026. https://doi.org/10.1007/s41060-025-00892-y

[13] O. Sener and V. Koltun, "Multi-task learning as multi-objective optimization," *arXiv preprint*, arXiv:1810.04650, 2018. https://doi.org/10.48550/arXiv.1810.04650

[14] S. Ruder, "An overview of multi-task learning in deep neural networks," *arXiv preprint*, arXiv:1706.05098, 2017. https://doi.org/10.48550/arXiv.1706.05098

[15] H. Matsuo and A. Khiat, "Early detection of reduced alertness using subsidiary behavior," SAE Technical Paper 2012-01-0737, 2012. https://doi.org/10.4271/2012-01-0737

[16] J. F. Wierwille and L. A. Ellsworth, "Evaluation of driver drowsiness by trained raters," *Accident Analysis & Prevention*, vol. 26, no. 5, pp. 571–581, 1994. https://doi.org/10.1016/0001-4575(94)90036-1

[17] Q. Han *et al.,* "A dense multi-pooling convolutional network for driving fatigue detection," *Scientific Reports*, vol. 15, 2025. https://doi.org/10.1038/s41598-025-99441-7

[18] H. Mohammadi, A. H. Alavi, and M. R. Daliri, "Driver drowsiness estimation using EEG signals with a dynamical encoder–decoder modeling framework," *Scientific Reports*, vol. 12, article 2483, 2022. https://doi.org/10.1038/s41598-022-06308-2

[19] M. S. Awais, N. Badruddin, and M. Drieberg, "PERCLOS threshold for drowsiness detection during real driving," *International Journal of Vehicular Technology*, 2017, article 4357356. https://doi.org/10.1155/2017/4357356

[20] A. Kapoor, W. Burleson, and R. Picard, "System and method for driver drowsiness detection using behavioral and sensor-based physiological measures," *Sensors*, vol. 23, no. 3, article 1292, 2023. https://doi.org/10.3390/s23031292

[21] National Academies of Sciences, Engineering, and Medicine, *LED Roadway Lighting: Impact on Driver Sleep Health and Alertness*. Washington, DC: The National Academies Press, 2021.

[22] H. Behbahani, N. Nadimi, H. Alenoori, and M. Sayadi, "Developing a new surrogate safety indicator based on motion equations," *Promet – Traffic & Transportation*, vol. 26, no. 5, pp. 451–460, 2014. https://doi.org/10.7307/ptt.v26i5.1388

[23] A. R. A. van der Horst and J. H. Hogema, "*Time-to-collision and collision avoidance systems*," University of Groningen, Groningen, Netherlands, Research Report, 1994.

[24] European New Car Assessment Programme, "*AEB Car-to-Car Systems Test Protocol*," Euro NCAP, 2022.

[25] T. D. Gillespie, *Fundamentals of Vehicle Dynamics*. Warrendale, PA: SAE International, 1992.

[26] R. Rajamani, *Vehicle Dynamics and Control*, 2nd ed. New York: Springer, 2012.

[27] U. Kiencke and L. Nielsen, *Automotive Control Systems: For Engine, Driveline, and Vehicle*, 2nd ed. Berlin: Springer, 2005.

[28] N. L. C. Thanh *et al.,* "Design and install anit-lock braking system (ABS) combined with regenerative braking system on 110cc honda lead motorcycle", *Proceedings of the 7th ATIGB International Scientific Conference,* 2022, pp. 79–84.

[29] N. L. C. Thanh, N. V. Dong, and N. V. Hai, "ABS technology solutions to increase braking performance for motorcycle", *The University of Danang - Journal of Science and Technology,* vol. 21, no. 5, pp.13-18, 2023. https://jst-ud.vn/jst-ud/article/view/8384