A Context-Aware Driver Assistance Framework Using Multimodal Attentiveness Monitoring and TTC-Based Risk Fusion

Tri Khoa Minh, Huy Pham Quang, An Nguyen Duc, Dung Doan Anh, Tung Nguyen Thanh*

School of Mechanical Engineering, Hanoi University of Science and Technology, Ha Noi, Vietnam

* *Corresponding author email: tung.nguyenthanh@hust.edu.vn*

Abstract

This study proposes a driver-monitoring and active-safety framework that integrates multimodal driver-state assessment with environmental risk evaluation. Driver attentiveness is estimated from visual and tactile cues, including eye behavior, head pose, hand posture, and steering-wheel touch interaction, while an external perception module detects frontal obstacles and estimates range using a vision-based deep learning approach. Based on the estimated range and relative closing velocity, Time-to-Collision (TTC) is calculated to represent collision urgency. A scoring-based fusion strategy combines the attentiveness score and TTC into a unified safety decision state that adjusts vehicle-control intervention according to both driver condition and environmental risk. The framework is implemented on a distributed embedded architecture and evaluated on a 1:10-scale vehicle platform under focused, distracted, and critical driving conditions. Experimental results show that incorporating TTC changes intervention behavior, with the largest descriptive difference observed under distracted driving and only a small additional effect under the critical condition. None of the paired Score-only versus Score–TTC comparisons reached statistical significance. Accordingly, the results are interpreted as proof-of-concept evidence that driver-state information and collision-risk information can be integrated within a unified closed-loop intervention framework, rather than as evidence of statistically established safety superiority. The proposed architecture provides a practical basis for further validation under more representative vehicle and traffic conditions.

Keywords: Distance estimation, driver state, safety, sensor, smart.

1. Introduction

Road safety remains a critical concern, with driver distraction and drowsiness contributing substantially to traffic accidents [1, 2]. Driver Monitoring Systems (DMS) estimate attentiveness from visual cues such as eye behavior, head pose, and hand activity [3-6], while multimodal approaches additionally incorporate tactile information to improve reliability [8-11]. In our previous work [12], eye behavior, hand interaction, head pose, and steering-wheel touch were fused into a scoring-based attentiveness index. However, that framework focused on driver-state assessment rather than environmentally aware vehicle intervention.

Active-safety systems commonly use Time-to-Collision (TTC) to represent collision urgency from range and relative closing velocity [13-15]. Driver monitoring and collision-risk assessment are generally treated as separate functions, although intervention should depend on both driver attentiveness and environmental urgency. Recent studies have advanced multimodal driver-state monitoring and TTC-based collision-risk assessment, but these two functions are still commonly investigated as separate decision processes; their integration within a low-cost embedded closed-loop intervention architecture remains comparatively underexplored [1, 10, 13]. Integrating these factors therefore enables a context-aware control strategy that can respond to both driver-related and collision-related risk.

Building on [12], the present study retains the eye-state, hand-posture, capacitive-touch, head-pose, and scoring-based attentiveness modules as the driver-monitoring front end, and introduces four extensions: (i) frontal-obstacle perception and TTC-based risk assessment; (ii) a Score–TTC fusion and control-scaling strategy; (iii) implementation on a distributed Jetson Nano–ESP32 closed-loop architecture; and (iv) experimental comparison of Score-only and Score–TTC intervention on a 1:10-scale vehicle under focused, distracted, and critical driving conditions. These additions extend the previous attentiveness-monitoring framework toward context-aware active intervention.

2. Methodology

The proposed framework integrates multimodal driver monitoring, frontal-obstacle perception, and embedded vehicle control. Driver-related inputs include eye state, hand–wheel interaction, capacitive touch, and head pose, while the environmental branch estimates

p-ISSN 3093-3285
e-ISSN 3093-3315
https://doi.org/10.51316/jst.xxx.ssad.xxxx.xx.x.x
Received: Jul 17, 2026; Revised: Aug 16, 2026;
Accepted: Aug 24, 2026; Online: Sep 8, 2026

Journal of Science and Technology — Smart Systems and Devices
Volume x, Issue x, Month Year, ppp–ppp

obstacle range and relative closing velocity for TTC computation.

The attentiveness score and TTC are fused into a context-dependent control state. The Jetson Nano performs perception and decision processing, while the ESP32 executes low-level vehicle control through the distributed UDP/serial architecture shown in Fig. 1. The attentiveness-monitoring branch is inherited from [12], whereas TTC integration and Score–TTC intervention constitute the functions evaluated in the present study.

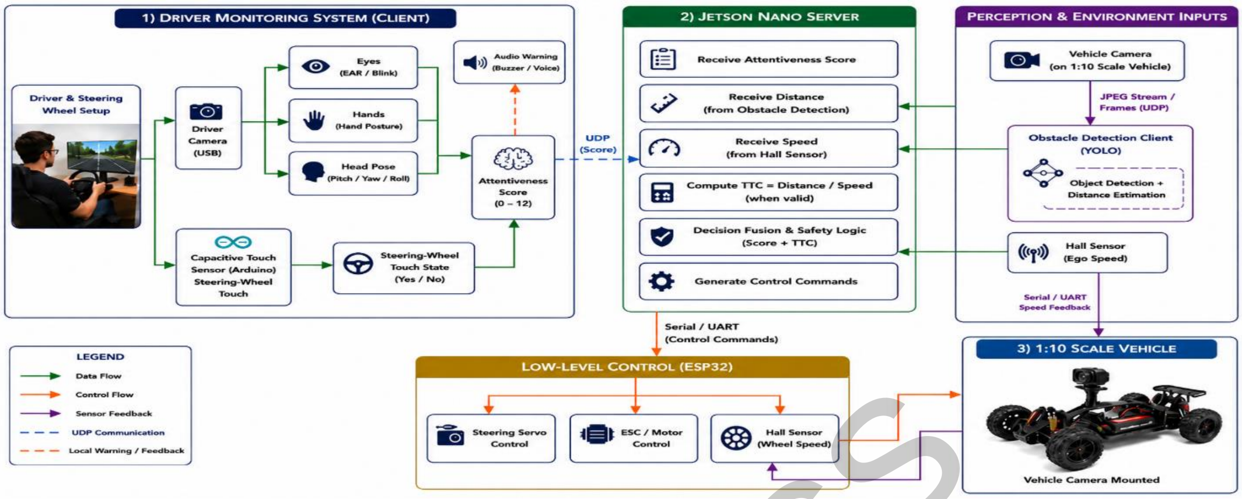

Fig. 1. Proposed system architecture

3. Multimodal Driver Attentiveness Monitoring

3.1. Drowsiness Detection

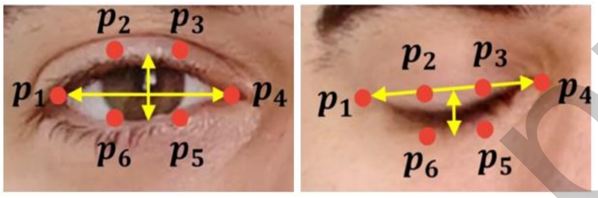

Fig. 2. Eye landmarks used for EAR calculation

$$EAR = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2 \times ||P_1 - P_4||} \quad (1)$$

Where P_1 and P_4 represent the horizontal eye corners, and P_2, P_3, P_5, P_6 represent the vertical eye landmarks. EAR [4] represents eyelid openness and decreases as the eye closes. In the experimental observations, mean EAR values were approximately 0.27–0.28 for open eyes, 0.20–0.26 for squinted eyes, and below 0.15 for closed eyes, as illustrated in Fig. 3. Based on these distributions, EAR > 0.26 represents an alert ocular state, 0.20 ≤ EAR ≤ 0.26 indicates reduced vigilance, and EAR < 0.20 represents increasing drowsiness-related risk. Values below 0.12 are associated with prolonged eye closure [4, 5].

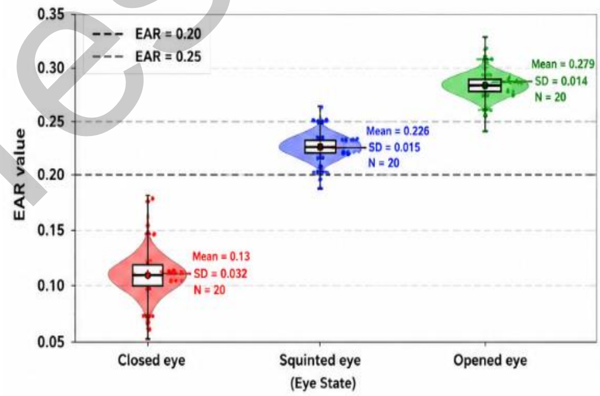

Fig. 3. EAR values for open, squinted, and closed eye states

EAR is evaluated together with eye-closure duration and blink behavior to distinguish normal blinking from sustained closure. The resulting ocular-state categories and scores are summarized in Table 1. These EAR and duration ranges are empirical categories used by the scoring framework and should not be interpreted as direct estimates of Karolinska Sleepiness Scale (KSS) levels, since KSS self-reports were not collected in the present experiment.

Table 1. Eye aspect ratio mapping

Observed State	**Average EAR**	**Duration (s)**	**Score**
Eyes wide open, normal blinking	EAR > 0.26	Blink < 1	0
Squinted eyes, reduced blink rate	0.20 ≤ EAR ≤ 0.26	Continuous ≥ 1	1
Slow blink, short closure	0.15 < EAR < 0.20	Continuous ≥ 1.5	2
Continuous eye closure	0.12 ≤ EAR ≤ 0.15	Continuous < 3	3
Prolonged eye closure	0.10 ≤ EAR < 0.12	Continuous 3–5	4
Very long eye closure	EAR < 0.10	Continuous > 5	5

Journal of Science and Technology — Smart Systems and Devices
Volume x, Issue x, Month Year, ppp–ppp

3.2. Steering-Wheel Hand–Touch Interaction Detection

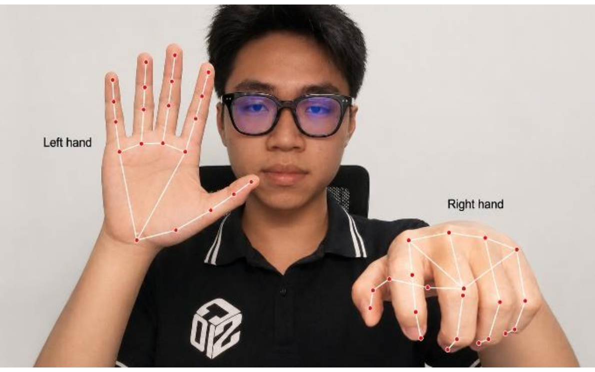

Fig. 4. Hand-posture detection in the driver-monitoring subsystem

Hand–wheel engagement is evaluated by combining vision-based hand-posture recognition with capacitive touch sensing. MediaPipe Hands extracts 21 hand landmarks from RGB images [6], while capacitive sensors verify physical contact with the steering wheel [7, 10]. For geometric analysis, the landmark coordinates are projected onto the image plane and scaled according to the frame dimensions before feature extraction.

Three geometric features are derived from the projected landmarks: palm length (wrist to middle fingertip), palm width (thumb base to little-finger base), and palm angle relative to the steering-wheel axis. Open-hand postures yielded mean values of 215.2, 242.1, and 2.82, respectively, whereas gripping postures yielded 146.2, 142.8, and 132.07. These feature distributions are used to distinguish open-hand and gripping-hand states, as illustrated in Fig. 5.

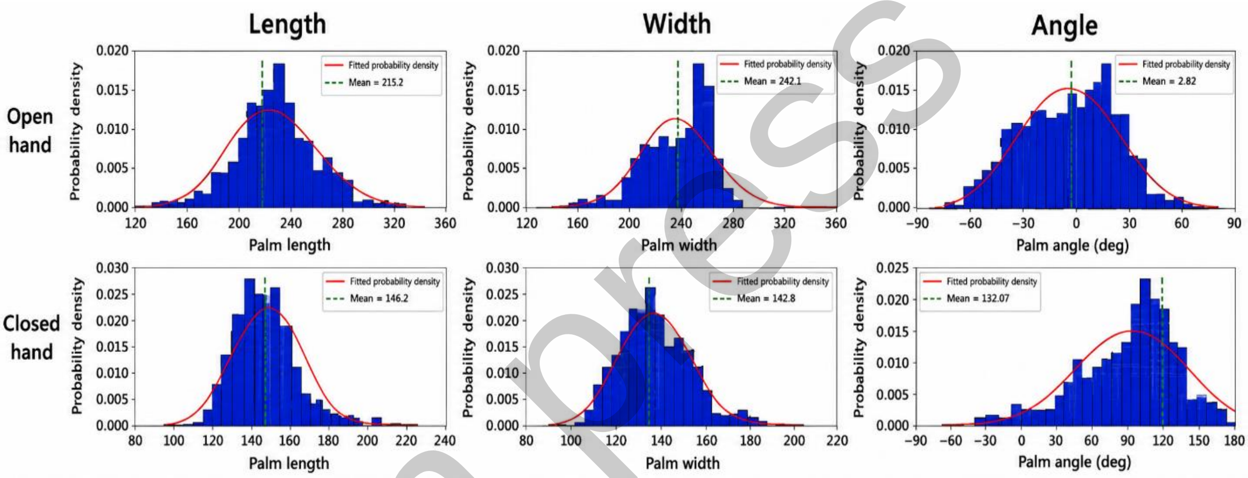

Fig. 5. Comparison of hand posture: open hand and closed hand on the steering wheel

Capacitive touch information is then combined with the visual hand state to distinguish secure grip, partial contact loss, and hand-release conditions. The duration of hand release is incorporated into the scoring process, with higher scores assigned to longer loss of hand–wheel contact. The resulting hand–touch states and scores are summarized in Table 2.

Table 2. Hand–touch mapping

Observed State	Duration (s)	Score
Gripping hand, touch	Any	0
Gripping hand, no touch	≥ 0.5	1
Open hand, touch	≥ 1	1
Short hand release	$1 < t \leq 2$	2
Medium hand release	$2 < t \leq 5$	3
Long hand release	$t > 5$	4

3.3. Head Pose Estimation

Head pose is represented by three Euler angles: pitch, yaw, and roll. Facial landmarks are extracted using MediaPipe Face Mesh, and the PnP algorithm is applied to estimate the three-dimensional head orientation [8]. The resulting pitch, yaw, and roll signals are filtered and normalized before a linear SVM classifies them into five states: forward, left, right, upward, and downward [9]. To avoid classifying brief natural movements such as mirror checking as distraction, a temporal constraint is applied so that a deviated orientation must persist for more than 2 s before being treated as sustained.

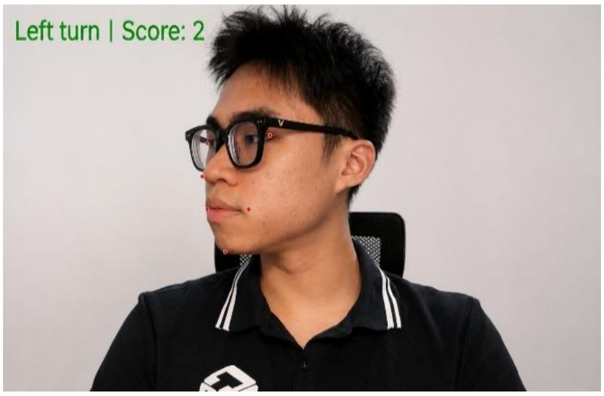

Fig. 6. Head-pose estimation in the driver-monitoring subsystem

The five orientation states are mapped to distraction scores according to direction and duration, as summarized in Table 3. Forward orientation is assigned a score of 0, while sustained lateral, upward, and downward deviations receive progressively higher scores. Prolonged downward orientation receives the highest score because of its association with fatigue and microsleep-related behavior [10]. The resulting head-pose score is subsequently integrated into the multimodal attentiveness index described in Section 3.4.

Table 3. Head pose mapping

Observed State	Duration (s)	Score
Forward facing	Any	0
Turning left or right	≤ 2	1
Turning left or right	> 2	2
Head up	≤ 2	2
Head up	> 2	3
Head down	≤ 2	2
Head down	> 2	4

3.4. Scoring-Based Evaluation Framework

The scoring-based attentiveness framework inherited from our previous work [12] combines eye behavior, hand–wheel interaction with capacitive touch, and head-pose estimation into a unified driver-state index. Instead of evaluating each modality separately, the proposed method converts visual, tactile, and temporal features into severity scores and aggregates them into a single attentiveness index.

Eye behavior is evaluated using EAR-based features to detect normal blinking, drowsiness, and prolonged eye closure [4, 5]. Hand–wheel engagement is assessed by combining hand landmark geometry and capacitive touch sensing to distinguish between secure grip, partial contact loss, and hands-off states [6, 7]. Head pose estimation contributes additional risk information by mapping pitch, yaw, and roll deviations into duration-based scores [8-10].

The scoring-based attentiveness framework inherited from [12] combines the eye, hand–touch, and head-pose scores into a unified driver-state index:

$$A = S_{eye} + S_{hand} + S_{head} \quad (2)$$

The corresponding modules were quantitatively evaluated in [12], with reported accuracies of 95.4% for drowsiness detection, 93.1% for hand detection, 98.2% for capacitive-touch detection, 94.7% for head-pose estimation, and 98.6% for overall attentiveness classification. These modules provide the attentiveness-score input used in the present Score–TTC study.

Unit weights and score thresholds are inherited from [12] to preserve comparability with the Score-only baseline. No modality-weight optimization or leave-one-modality-out ablation was performed; equal weighting is therefore treated as an inherited empirical design choice rather than an optimized fusion rule. The resulting score is mapped to the attention levels in Table 4.

Table 4. Scoring framework

Total Score	Attention level	Behavioral Description & Recommended Action
0 – 2	Highly attentive	Safe condition, no warning required
3 – 5	Alert	Mild signs present, no warning necessary
6 – 8	Distracted	Issue light warning
9 – 11	Dangerous	Issue critical warning
≥12	Extremely Dangerous	Stop the vehicle, activate emergency ADAS

4. Active Safety Perception, Decision, and Control Framework

4.1. Obstacle Perception, Speed Measurement, and TTC Evaluation

Fig. 7. Frontal-obstacle detection and monocular distance estimation using the camera mounted on the 1:10-scale vehicle

In this study, a monocular camera mounted at the front of the vehicle is used to detect frontal obstacles. Frontal-obstacle detection is performed using a YOLO-based model. The detector was evaluated on 465 annotated images containing 465 object instances. It achieved a precision of 0.996, a recall of 0.603, an mAP@0.50 of 0.991, and an mAP@0.50:0.95 of 0.743. The high precision and mAP@0.50 indicate accurate detections for successfully identified targets, whereas the lower recall shows that missed detections remain a limitation of the current perception module. From the detected objects, the nearest obstacle in the forward driving direction is selected as the primary target for risk evaluation.

The driver-monitoring modules do not directly contribute to TTC estimation; eye state, hand–touch interaction, and head pose influence the intervention through the attentiveness-score branch, whereas TTC is determined by obstacle detection, monocular range estimation, and vehicle-speed measurement.

The distance to the detected obstacle is estimated using the pinhole camera model:

$$R = \frac{f.H}{h} \tag{3}$$

where f denotes the camera focal length, H is the real-world height of the detected object, and $h$is the height of the corresponding bounding box in the image.

The camera parameters used in the pinhole model were calibrated before the vehicle experiments, while corresponds to the known physical height of the reference target used in the test setup. The same camera configuration and reference-target definition were maintained throughout the experiments. However, an independent ground-truth range-error dataset was not retained; therefore, the present study does not claim quantitatively validated absolute accuracy of the monocular range estimate.

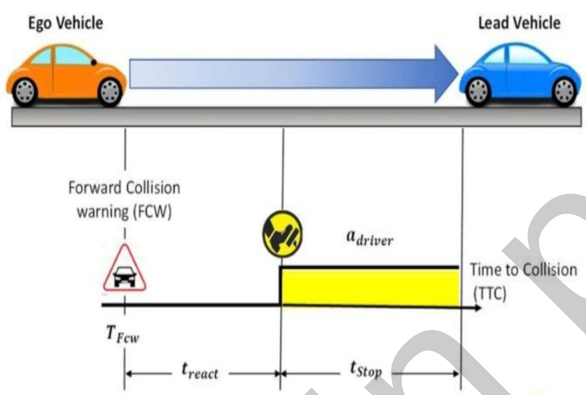

Fig. 8. Illustration of Time-to-Collision (TTC) concept

Because TTC depends directly on the estimated range R, an uncertainty ΔR produces, for fixed v_{cl}, an approximate TTC error of:

$$\Delta TTC \approx \frac{\Delta R}{v_{cl}} \tag{4}$$

Systematic ground-truth characterization of the monocular range estimator across the operating distance remains necessary for future validation.

Vehicle speed is measured using a Hall-effect sensor mounted on the drivetrain. The sensor generates pulse signals corresponding to wheel rotation, which are processed to calculate the vehicle's linear velocity in real time. The speed signal is filtered to reduce measurement noise before synchronization with the range estimate.

Based on the estimated longitudinal range and relative closing velocity, TTC is computed as:

$$TTC = \frac{R}{v_{cl}} \tag{5}$$

$$v_{cl} = v_{ego} - v_{target} \tag{6}$$

Where denotes the longitudinal range between the ego vehicle and the target, and v_{cl} is the relative closing velocity. For a target moving in the same longitudinal direction, $v_{cl} = v_{ego} - v_{target}$. TTC is evaluated only when $v_{cl} > 0$, indicating that the range between the ego vehicle and the target is decreasing.

In the present experiments, the frontal obstacle was stationary; therefore, $v_{target} = 0$ and $v_{cl} = v_{ego}$. Under this condition, the general TTC formulation reduces to $TTC = \frac{R}{v_{cl}}$, which is equivalent to the expression implemented in the experimental controller.

In practical implementation, TTC is considered invalid when no obstacle is detected or when vehicle speed approaches zero, preventing unstable or meaningless calculations. When TTC is invalid because no valid obstacle is detected or the relative closing velocity is non-positive, the TTC branch is excluded from the fusion decision, and the controller temporarily falls back to the attentiveness-score-based mode until a valid TTC estimate becomes available. The computed TTC value is then categorized into different risk levels, ranging from safe driving conditions to imminent collision scenarios. No dedicated temporal filter was applied to the monocular range or TTC estimates in the present implementation. Temporal smoothing of range and TTC, for example using Kalman or alpha–beta filtering, should be investigated in future work to reduce frame-to-frame fluctuations and improve the stability of collision-risk estimation.

4.2. Fusion-Based Safety Decision Strategy

The decision module independently evaluates the attentiveness score and TTC and fuses their risk levels into the control modes defined in Table 5. The attentiveness branch represents driver-related risk, whereas TTC represents environmental collision urgency [13].

Table 5. Mapping between risk conditions and control scaling

Condition	**Mode**	**Control Scale**
Score ≤ 6 and TTC ≥ 3 s	Normal	1.0
7 ≤ Score ≤ 9 or 2 ≤ TTC < 3 s	Reduced	0.7
10 ≤ Score < 12 or 1 ≤ TTC < 2 s	Limited	0.3
Score ≥ 12, TTC < 1 s, or both risk branches ≥ Limited	Safe	0.0

The TTC bands and control-scaling factors are empirical parameters selected for the 1:10-scale proof-of-concept platform. No dynamic-similitude law was applied; therefore, the reported thresholds, control gains, and model-scale distances should not be quantitatively extrapolated to full-size vehicles and require independent calibration for real-vehicle application.

Normal, Reduced, and Limited modes progressively restrict longitudinal control as risk increases. Safe mode sets the control scale to 0.0 when either the attentiveness score reaches or exceeds 12, TTC falls below 1 s, or both risk branches simultaneously reach at least the Limited level.

5. Experimental Setup

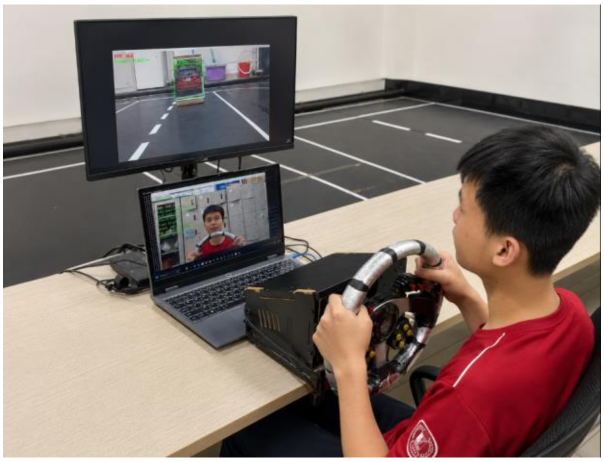

Fig. 9. Experimental setup showing the participant, driver-monitoring camera, steering-wheel interface, forward-facing vehicle camera, and 1:10-scale vehicle platform

A final-clearance experiment was conducted on the developed 1:10-scale vehicle platform to evaluate the intervention behavior of the proposed driver-monitoring and active-safety control system. The vehicle was equipped with a front-facing camera, a driver-monitoring camera, capacitive touch sensors on the steering wheel, a Jetson Nano, and an ESP32. The Jetson Nano processed driver-state recognition, obstacle detection, distance estimation, TTC calculation, and decision-making, while the ESP32 executed throttle, braking, and steering commands.

During each trial, the participant operated the 1:10-scale vehicle through the steering-wheel control interface in the laboratory. A driver-facing camera was positioned in front of the participant to capture eye behavior, head pose, and hand–wheel interaction for the driver-monitoring subsystem. A separate forward-facing camera was mounted on the 1:10-scale vehicle and was used exclusively for frontal-obstacle detection and monocular distance estimation. The Jetson Nano processed the driver-monitoring, obstacle-perception, and Score–TTC decision functions, while the ESP32 executed the corresponding low-level steering and longitudinal-control commands.

The inherited driver-monitoring pipeline reported in our previous work [12] has a measured processing latency of 0.28 s. In the present distributed implementation, the low-level vehicle-control loop operates at approximately 50 Hz, corresponding to a 20 ms control-update period. The Jetson Nano performs the high-level perception and decision tasks, whereas the ESP32 executes steering and ESC commands. A separate end-to-end measurement covering the complete perception-to-actuation chain was not recorded in the present experiment.

The experiment was performed in a controlled laboratory environment with a fixed obstacle placed in front of the vehicle. In each trial, the participant drove the vehicle toward the obstacle, while the system applied the corresponding intervention strategy. After the vehicle came to rest, the remaining longitudinal gap between the front of the vehicle and the obstacle was recorded as the final clearance distance, d_{clear}. This quantity is distinct from braking distance. A smaller positive d_{clear} indicates that the vehicle stopped closer to the obstacle and therefore represents a later or less conservative intervention; it should not, by itself, be interpreted as evidence of improved safety. Because a target safe-clearance band was not predefined in the original experimental protocol, the final-clearance values are interpreted descriptively as indicators of intervention timing rather than as direct measures of safety performance.

The dataset included 20 participants (average age = 20, all male with valid driving licenses), with each record containing seven test conditions: Familiarization Trial (FAM), Focused with Score-based Control (FOC-S), Focused with Score–TTC Control (FOC-ST), Distracted with Score-based Control (DIS-S), Distracted with Score–TTC Control (DIS-ST), Critical with Score-based Control (CRT-S), and Critical with Score–TTC Control (CRT-ST). The 20-participant sample was selected for this proof-of-concept laboratory study within the available experimental protocol rather than through an a priori power analysis. Accordingly, the statistical findings should be interpreted cautiously, and larger and more diverse cohorts are required for confirmatory validation. These conditions were used to compare the effectiveness of score-based intervention and score–TTC intervention under different driver states.

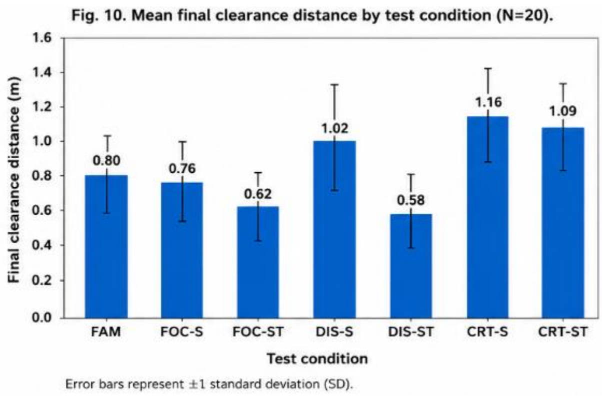

Fig. 10. Mean final clearance distance by test condition (N=20)

All experiments were conducted in a controlled laboratory environment using the 1:10-scale vehicle platform, with no on-road driving involved. Before testing, each participant was informed of the purpose of the study, experimental procedure, types of data to be collected, and their intended research use, and voluntarily agreed to participate. The collected data was limited to driving-related behavioral and control-response information generated during the experimental tasks and was retained in the laboratory's internal research dataset. Participant briefing, data collection, and experimental procedures followed the applicable ethical principles and laboratory procedures for human-related research data collection. No formal institutional ethics approval number was issued for this study; therefore, no ethics-approval identifier is reported.

6. Results and Discussion

The experiments examine how TTC modifies the intervention behavior of the attentiveness-based controller. Across the three driver states, the Score–TTC strategy produced lower mean final-clearance values than Score-only control. A lower final clearance indicates that the vehicle stopped closer to the obstacle and reflects a later or less conservative intervention; it is not interpreted independently as evidence of improved safety.

No obstacle contact occurred in any valid experimental trial, corresponding to a collision rate of 0%. The minimum observed final clearances were 0.309 m for FOC-S, 0.369 m for FOC-ST, 0.341 m for DIS-S, 0.322 m for DIS-ST, 0.425 m for CRT-S, and 0.523 m for CRT-ST. The overall minimum observed clearance was 0.309 m. These values characterize the outcomes of the present experiments and are not interpreted as predefined safety-clearance thresholds.

In the focused condition, the mean final clearance changed from 0.764 m in FOC-S to 0.619 m in FOC-ST, corresponding to an 18.9% relative difference. Because the attentiveness score remained low, the lower clearance is consistent with TTC delaying intervention until environmental collision urgency increased. This result is interpreted as a change in intervention timing rather than as an independent safety improvement.

The largest descriptive difference was observed in the distracted condition, where the mean final clearance changed from 1.037 m in DIS-S to 0.572 m in DIS-ST, corresponding to a 44.9% relative difference. The Score–TTC strategy incorporates obstacle proximity and vehicle motion in addition to the driver-state score, which modifies the timing of intervention as collision urgency increases.

In the critical condition, the mean final clearance changed from 1.151 m in CRT-S to 1.094 m in CRT-ST, corresponding to a 4.9% relative difference. Because the attentiveness score was already sufficiently high to trigger strong intervention, the additional TTC input produced only a small change in the resulting control behavior.

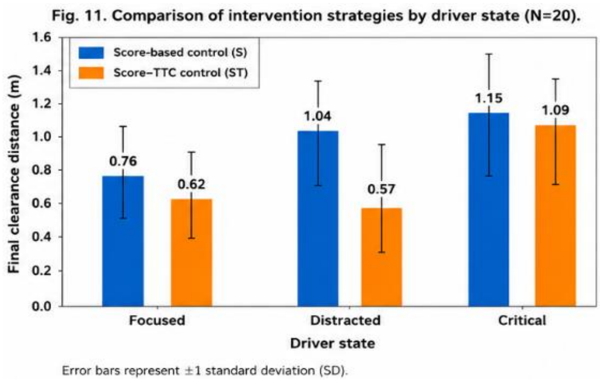

Fig. 11. Comparison of intervention strategies by driver state (N=20)

Because the Score-only and Score–TTC conditions were evaluated within the same 20-participant cohort, paired t-tests were performed for each driver state using the available complete paired records. For the focused condition, the mean paired difference (Score-only minus Score–TTC) was 0.145 m (95% CI: −0.199 to 0.483 m, p = 0.365, Cohen's $d_z = 0.32$). For the distracted condition, the mean paired difference was 0.465 m (95% CI: −0.048 to 0.975 m, p = 0.068, $d_z = 0.84$). For the critical condition, the mean paired difference was 0.057 m (95% CI: −0.487 to 0.586 m, p = 0.834, $d_z = 0.08$). None of the three paired comparisons reached the conventional $p < 0.05$ significance level. Accordingly, the percentage differences reported above are interpreted descriptively rather than as statistically significant improvements.

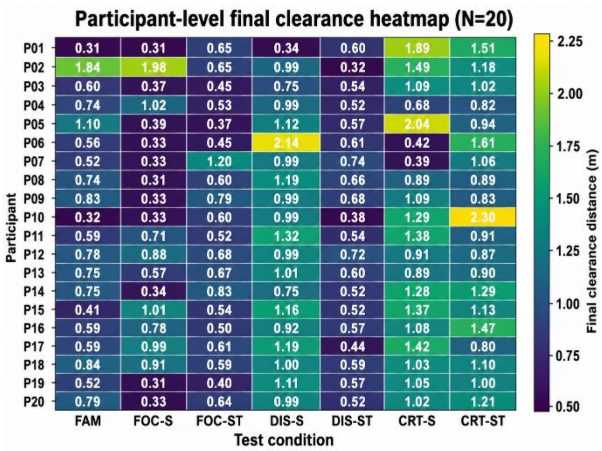

Fig. 12. Distribution of final clearance distance across the experimental conditions

The distribution of final clearance is shown on Fig. 12. A narrower descriptive spread is observed for DIS-ST than for DIS-S in the available dataset. However, because no dedicated repeatability or variance analysis was performed, this observation is interpreted only as lower observed inter-participant dispersion and not as evidence of improved repeatability or control stability.

External validity is limited by the controlled indoor setting, stationary obstacle, 1:10-scale platform, and narrow participant demographic. TTC was evaluated only for a stationary frontal obstacle under approximately constant relative velocity; moving-target scenarios were not tested. The sample comprised 20 male drivers with a mean age of 20 years, limiting generalization across age, sex, and eyewear. Full-scale testing with a more diverse participant population is required for further validation.

7. Conclusion

This study extends the multimodal attentiveness framework in [12] by incorporating TTC-based environmental risk into a closed-loop vehicle-control strategy. On the 1:10-scale platform, TTC produced the largest descriptive change in intervention timing under distracted driving, while its additional effect was small when the attentiveness score was already critical. None of the paired comparisons reached statistical significance; therefore, the results support proof-of-concept feasibility rather than superior safety performance. Current limitations include monocular range uncertainty, stationary-obstacle testing, model-scale dynamics, incomplete end-to-end timing characterization, and the narrow participant cohort. Future work will address calibrated range validation, moving targets, TTC-only comparison, full-system timing, and larger full-scale studies.

Acknowledgments

This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2025-PC-046.

References

[1] E. Michelaraki, C. Katrakazas, S. Kaiser, T. Brijs, and G. Yannis, Real-time monitoring of driver distraction: State-of-the-art and future insights, Accident Analysis & Prevention, vol. 192, Nov. 2023, Art. no. 107241. https://doi.org/10.1016/j.aap.2023.107241

[2] World Health Organization, Global status report on road safety 2018. Geneva, Switzerland: World Health Organization, 2018. [Online]. Available from the WHO publication repository.

[3] A. Koesdwiady, R. Soua, F. Karray, and M. S. Kamel, Recent trends in driver safety monitoring systems: State of the art and challenges, IEEE Transactions on Vehicular Technology, vol. 66, iss. 6, pp. 4550–4563, Jun. 2017 https://doi.org/10.1109/TVT.2016.2631604

[4] T. Soukupová and J. Čech, Real-time eye blink detection using facial landmarks, in Proc. 21st Computer Vision Winter Workshop (CVWW), Rimske Toplice, Slovenia, pp. 1–8, Feb. 2016.

[5] A. Picot, S. Charbonnier, and A. Caplier, On-line detection of drowsiness using brain and visual information, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, vol. 42, iss. 3, pp. 764–775, May 2012. https://doi.org/10.1109/TSMCA.2011.2164242

[6] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.-L. Chang, and M. Grundmann, MediaPipe Hands: On-device real-time hand tracking, arXiv preprint arXiv:2006.10214, Jun. 2020. https://doi.org/10.48550/arXiv.2006.10214

[7] W. Li, J. Huang, G. Xie, F. Karray, and R. Li, A survey on vision-based driver distraction analysis, Journal of Systems Architecture, vol. 121, Art. no. 102319, Dec. 2021. https://doi.org/10.1016/j.sysarc.2021.102319

[8] E. Murphy-Chutorian and M. M. Trivedi, Head pose estimation in computer vision: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, iss. 4, pp. 607–626, May 2009. https://doi.org/10.1109/TPAMI.2008.106

[9] C. Cortes and V. Vapnik, Support-vector networks, Machine Learning, vol. 20, pp. 273–297, Sep. 1995. https://doi.org/10.1007/BF00994018

[10] M. S. Al-Quraishi, S. S. Azhar Ali, M. Al-Qurishi, T. B. Tang, and S. Elferik, Technologies for detecting and monitoring drivers' states: A systematic review, Heliyon, vol. 10, iss. 20, Oct. 2024, Art. no. e39592. https://doi.org/10.1016/j.heliyon.2024.e39592

[11] Y. Ma, V. Sanchez, S. Nikan, D. Upadhyay, B. Atote, and T. Guha, Robust multiview multimodal driver monitoring system using masked multi-head self-attention, in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 2617–2625, Aug. 2023. https://doi.org/10.1109/CVPRW59228.2023.00260

[12] K. M. Tri, P. Q. Huy, N. D. An, D. A. Dung, and N. T. Tung, A multimodal driver attentiveness monitoring framework via scoring-based fusion of visual and tactile cues, in *Proc. SETSM*, 2026, accepted for publication.

[13] D. R. Sarkar, K. R. Rao, and N. Chatterjee, A review of surrogate safety measures on road safety at unsignalized intersections in developing countries, Accident Analysis & Preventio*n*, vol. 195, Feb. 2024, Art. no. 107380. https://doi.org/10.1016/j.aap.2023.107380

[14] K. Vogel, A comparison of headway and time to collision as safety indicators, Accident Analysis & Prevention, vol. 35, iss. 3, pp. 427–433, May 2003. https://doi.org/10.1016/S0001-4575(02)00022-2

[15] J. B. Cicchino, Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates, Accident Analysis & Prevention, vol. 99, part. A, pp. 142–152, Feb. 2017. https://doi.org/10.1016/j.aap.2016.11.009

[16] Y. Yuan, Y. Lu, and Q. Wang, Adaptive forward vehicle collision warning based on driving behavior, Neurocomputing, vol. 408, pp. 64–71, Sep. 2020. https://doi.org/10.1016/j.neucom.2019.11.024

[17] L. Masello, G. Castignani, B. Sheehan, F. Murphy, and K. McDonnell, On the road safety benefits of advanced driver assistance systems in different driving contexts, Transportation Research Interdisciplinary Perspectives, vol. 15, Sep. 2020, Art. no. 100670. https://doi.org/10.1016/j.trip.2022.100670