LETTER

Real-Time Safety Driving Advisory System Utilizing a Vision-Based Driving Monitoring Sensor

Masahiro TADA[†a], *Member and* **Masayuki NISHIDA**[††], *Nonmember*

SUMMARY   In this study, we use a vision-based driving monitoring sensor to track drivers' visual scanning behavior, a key factor for preventing traffic accidents. Our system evaluates driver's behaviors by referencing the safety knowledge of professional driving instructors, and provides real-time voice-guided safety advice to encourage safer driving. Our system's evaluation of safe driving behaviors matched the instructor's evaluation with accuracy over 80%.
key words: driver monitoring, safety driving advisory system

1. Introduction

Recent advancements in the field of Advanced Driver Assistance Systems (ADAS), including the Advanced Emergency Braking System (AEBS), have garnered significant attention [1]. Examining the types of traffic accidents in Japan, rear-end collisions account for 30.5%, the highest percentage, and head-on collisions at intersections follow at 25.9%, making it the second most common accident type. Although the spread of AEBS is expected to reduce the number of rear-end collisions, detecting oncoming vehicles from crossroads at intersections with poor visibility using onboard sensors is difficult. Consequently, the preventive effects of AEBS against head-on collisions are believed to be limited.

According to the National Police Agency's statistics, over 70% of all traffic accidents in Japan stem from driver errors. The most frequent cause of traffic accidents is the driver's insufficient visual scanning behavior to ensure surrounding safety. Furthermore, statistics of traffic fatalities in Japan in 2022 reveals that pedestrians constituted the highest proportion, accounting for 36.6% of the deaths. Since detecting pedestrians using onboard vehicle sensors is more challenging compared to vehicles, these underscore the need for mitigating driver errors especially visual scanning errors in tandem with augmentations in vehicular developments such as ADAS to ensure a comprehensive reduction in traffic accidents.

To achieve this, in this paper, we propose a real-time safety driving advisory system using vision-based driving monitoring sensor to encourage drivers to behave safely at potentially dangerous spots where traffic risk increases.

Here, in this paper, we defined intersections where probability of collisions, both with vehicles and pedestrians, escalates as potentially dangerous spots, and defensive driving behaviors to prevent traffic accidents at potentially dangerous spots as safe driving behaviors.

We use vision-based driving monitoring sensor which employs deep-learning technology to track the driver's facial orientation. Our system evaluates driver's safe driving behaviors using the facial orientation data by referencing the knowledge of professional driving instructors from the viewpoint of active safety for mitigating traffic accident risks. If a driver behaves riskily (e.g. a driver approaching a blind intersection without adequate visual scanning), our system provides voice-guided safety advice based on evaluation results and encourages him/her to drive safer.

2. Related Work

For driving assistance, various systems focus on behaviors like driver drowsiness and distraction detection [2]–[5]. These systems primarily detect abnormal conditions such as drowsiness or inattention (e.g., looking away, using a phone while driving). Nonetheless, drivers can also engage in risky behavior under normal conditions, such as failing to scan surroundings at a blind intersection. Identifying such potentially hazardous behaviors is crucial for assessing safe driving behaviors, particularly in how drivers scan their environment to mitigate traffic risks.

Recent studies have attempted to detect cognitive distraction in drivers, where attention deviates despite eyes facing forward, under simulator [6] and real-world traffic conditions [7]. In prior research on real-time safety driving advisory systems, Tanaka et al. proposed a system using a small robot-type agent that verbally notifies drivers of the presence of potential hazard at pre-registered locations such as stop intersections and communicates what safety scans should be performed there through the robot's gestures before entering the intersection [8], [9]. Chen et al. have proposed a system that evaluates the appropriateness of passing speed and distance from parked vehicles when overtaking parked cars, providing feedback to the driver through the color of LEDs [10]. However, these studies neither aim to measure and evaluate a driver's visual scanning behaviors in real-time, nor do they focus on the real-time identification of potentially hazardous behaviors, nor do they provide safety driving advice in real-time to encourage him/her to drive safer.

Tada et al. [11] developed an automatic safe driving behaviors evaluation system using wearable sensors. This system measures the driver's face orientation with a wearable motion sensor to detect scanning behaviors. Combining these measurements with GPS data and driving instructor insights, it automatically assesses a driver's safe-driving behaviors, focusing on visual scanning behaviors at intersections to detect potentially hazardous behaviors. However, the requirement for drivers to wear motion sensors makes widespread application challenging.

3. Safety Driving Advisory System

3.1 System Components

In our previous work [11], we explored the correlation between face orientation data and eye-camera data, verifying that face orientation effectively indicates visual scanning behaviors at intersections. Our system comprises a vision-based driving monitoring sensor (130 mm × 60 mm × 130 mm), a gyro sensor (37 mm × 46 mm × 12 mm), a GPS receiver, and a laptop PC, as shown in Fig. 1. The vision-based driver monitoring sensor (DriveKarte®, OMRON SOCIAL SOLUTIONS, Co. Ltd.) employs an infrared camera to monitor the driver's face and employs deep learning technology to determine the driver's face orientation. This sensor, mounted on the vehicle's dashboard, sends data to a laptop PC at 30 Hz sampling rate. A gyro sensor, measuring vehicle movements for right and left turns at intersections, transmits data at 25 Hz to the PC via Bluetooth. The laptop PC records these data and aligns it with GPS time.

The system monitors GPS locations at 1 Hz, triggering a safe driving behaviors evaluation when the vehicle enters predetermined intersections identified as potentially dangerous by a professional driving instructor. By providing safety driving advice based on evaluation criteria described in Sect. 3.2, our system encourage safer driving.
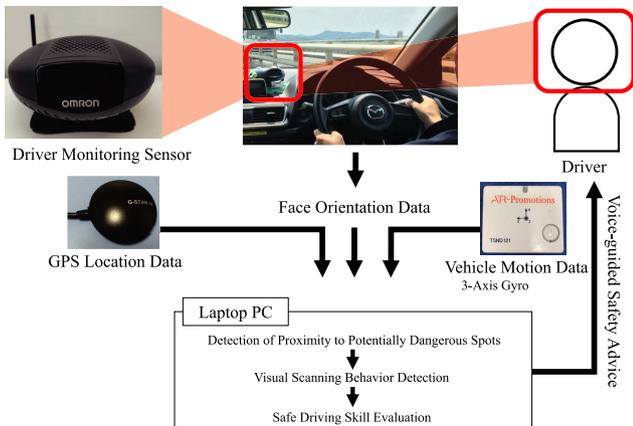

Fig. 1 System overview.

3.2 Criteria for Evaluating Safe Driving Behaviors

Based on interviews with three professional driving instructors, Tada et al. reported that instructors typically evaluate safe driving behaviors based on criteria as follows [11].

Visual Scanning Behavior To mitigate collision risks at intersections, it's crucial for drivers to scan their surroundings thoroughly, including significant head rotation. Adequate visuall scanning ensures the gathering of enough information to ascertain safety. Context determines scanning appropriateness: extensive head rotation is beneficial at intersections to avoid collisions with cyclists or pedestrians, yet it might be distracting and dangerous on expressways. Driving instructors evaluate scanning behavior by its direction, face orientation angle, sufficient duration for information gathering, and appropriate timing (for instance, scanning after crossing an intersection is ineffective for collision prevention).

Vehicle Speed Appropriate speed maintenance is vital for safe driving. Early detection of vehicles, cyclists, or pedestrians is less effective if a driver's excessive speed hinders avoiding them. Slower speeds near potential hazards allow for more thorough scanning.

From the interview findings, Tada et al. [11] created a "minimum" safe-driving behavior checklist for potentially dangerous spots, focusing on: (1) scanning direction, (2) number of scans, (3) duration of each scan to ensure hazard detection, (4) scan timing, and (5) driving speed. Since various potentially dangerous spots necessitate different safety measures, we customized checklists for each, guided by driving instructors' expertise, as presented in Table 1. Labels 'A' to 'I' in Table 1 correspond to those in Fig. 3. The parameters θ_{deg}, θ_t and θ_v in Table 1 denote the threshold values for the maximum angle, duration time of scanning behavior and driving speed respectively, used to determine whether the observed driving behavior satisfies the respective evaluation item. Appropriate timing for each scanning behavior, classified as 'before entering spot' and 'during turning spot' in Table 1 is determined using GPS location data and vehicle movement data (gyro sensor data).

3.3 Procedure for Evaluating Safe Driving Behaviors

Our system's process for evaluating safe driving behaviors is as follows.

**Step1** Evaluation begins when a vehicle nears predetermined, instructor-identified potentially dangerous spots, as tracked via GPS.

Step2 Detect driver's visual scanning behavior through sequential facial orientation data from a vision-based sensor. To focus on meaningful scanning behaviors and disregard minor face movements, we consider only face orientation data with an angular velocity exceeding

Table 1   Checklists of minimum safe-driving.

Spots	Evaluation Items			
A	Scan left side-mirror before entering spot $\theta_{deg} \geq 10\text{deg}, \theta_t : \text{N.D.}$	Scan left-rear side before entering spot $\theta_{deg} \geq 50\text{deg}, \theta_t \geq 2.0\text{s}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$	–
B	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.5\text{s}$	Scan right-rear side during turning spot $\theta_{deg} \geq 50\text{deg}, \theta_t \geq 1.5\text{s}$	–	–
C	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.5\text{s}$	Scan right-rear side during turning spot $\theta_{deg} \geq 50\text{deg}, \theta_t \geq 1.5\text{s}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$	–
D	Scan left side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$	–
E	Scan left side before entering spot $\theta_{deg} \geq 35\text{deg}, \theta_t \geq 1.5\text{s}$	Scan right side before entering spot $\theta_{deg} \geq 35\text{deg}, \theta_t \geq 1.5\text{s}$	Scan left side during turning spot $\theta_{deg} \geq 30\text{deg}, \theta_t : \text{N.D.}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$
F	Scan left side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$	–
G	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.5\text{s}$	Scan right-rear side during turning spot $\theta_{deg} \geq 50\text{deg}, \theta_t \geq 1.5\text{s}$	–	–
H	Scan left side before entering spot $\theta_{deg} \geq 35\text{deg}, \theta_t \geq 1.5\text{s}$	Scan right side before entering spot $\theta_{deg} \geq 35\text{deg}, \theta_t \geq 1.5\text{s}$	Scan left side during turning spot $\theta_{deg} \geq 30\text{deg}, \theta_t : \text{N.D.}$	Stop at the spot $\theta_v \leq 0\text{km/h}$
I	Scan left side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	Scan right side before entering spot $\theta_{deg} \geq 30\text{deg}, \theta_t \geq 1.0\text{s}$	decelerate to 10km/h $\theta_v \leq 10.0\text{km/h}$	–

'N.D.' indicates that threshold values are not defined for a specific evaluation item.
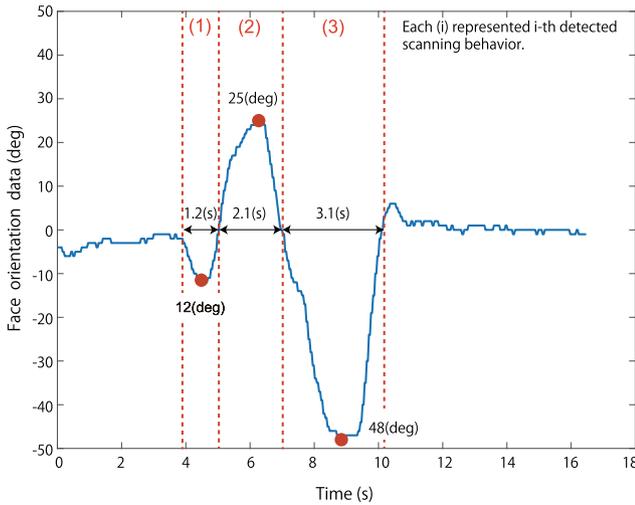
Fig. 2   Scanning behavior detection from head orientation data.

20 deg/s [11] as indicative of visual scanning behavior.

Step3 Calculate the face orientation angle d_i, duration time t_i of i-th detected visual scanning behavior as shown in Fig. 2, and represented as a feature vector $x_i = (d_i, t_i)$. Here, the driver monitoring sensor outputs positive values when the driver's face orientation is towards the left and negative values when oriented to the right. As mentioned in Step 2, face movements with an angular velocity below the threshold (20 deg/s) are not considered as visual scanning behaviors. Additionally, a change in the sign of the facial orientation angle is in-

Table 2   Scoring method of vehicle speed.

existence of a stop sign	compliance rate of vehicle speed	$score_v$
with a stop sign	$v = 0$	1.0
	$v > 0$	0.5
without a stop sign	$v \leq \theta_v$	1.0
	$\theta_v < v \leq \theta_v + 5.0$	0.9
	$\theta_v + 5.0 < v \leq \theta_v + 10.0$	0.7
	$\theta_v + 10.0 < v$	0

terpreted as the initiation of a different visual scanning behavior. In the example shown in Fig. 2, three visual scanning behaviors were detected: one scanning to the left side and two to the right.

Step4 Calculate vehicle minimum speed v at the spot from GPS location data, and score vehicle speed as shown in Table 2.

Step5 The driver's visual scanning behaviors are evaluated by matching observed behaviors (Steps 2–3) against pre-set evaluation items for each spot as shown in Table 1. Consider determining whether the evaluation item n "Scan right side before entering spot" with parameters $\theta_n = (\theta_{deg,n}, \theta_{t,n})$ at intersection s is fulfilled. Let $X_s = \{x_1, \ldots, x_M\}$ be the set of feature vectors for all visual scanning behaviors at intersection s that meet the criteria of direction (i.e. right) and timing (i.e. before entering spot) for the evaluation item n. The extent to which each $x_i = (d_i, t_i) \in X_s$ satisfies the evaluation item n is quantified based on the method shown in Table 3 and the highest score among these is assigned as the $score_{scan,n}$ for the evaluation item n.

Table 3 Scoring method of scanning behavior.

compliance rate of scanning angle	compliance rate of scanning time	voice-guided advice	$score_{scan}$
$\frac{|d_i|}{\theta_{deg,n}} \geq 1.0$	$\frac{d_t}{\theta_{t,n}} \geq 1.0$	—	5
$1.0 > \frac{|d_i|}{\theta_{deg,n}} \geq 0.5$	$\frac{d_t}{\theta_{t,n}} \geq 1.0$	The angle of your scanning to the left (right) side is slightly shallow	4
$1.0 > \frac{|d_i|}{\theta_{deg,n}} \geq 0.5$	$\frac{d_t}{\theta_{t,n}} < 1.0$	The angle of your scanning to the left (right) side is slightly shallow and scanning time is insufficient.	3
$0.5 > \frac{|d_i|}{\theta_{deg,n}} > 0$	$\frac{d_t}{\theta_{t,n}} \geq 1.0$	The angle of your scanning to the left (right) side is shallow.	3
$0.5 > \frac{|d_i|}{\theta_{deg,n}} > 0$	$\frac{d_t}{\theta_{t,n}} < 1.0$	The angle of your scanning to the left (right) side is shallow and scanning time is insufficient.	2
no scanning behavior	no scanning behavior	Scanning to the left (right) side is not confirmed.	0

Table 4 Grading criteria based on total score.

	grade
$score_{total} \geq 4.25$	A
$4.25 > score_{total} \geq 3.5$	B
$3.5 > score_{total} \geq 2.5$	C
$2.5 > score_{total} \geq 1.5$	D
$score_{total} < 1.5$	E

Step6 Calculate total score of driving behaviors at the spot using Eq. (1). Here, $score_{scan,n}$, N, $score_v$ represent scanning score for evaluation item n at the spot calculated in Step 5, the number of evaluation items related to scanning behavior at the spot, and the vehicle speed score calculated in Step 4, respectively. The driver's behaviors at each spot is graded from A (Excellent) to E (Worst), based on the total score as shown in Table 4. If driving behaviors at the spot are graded as 'A', the system acknowledges this with a "good driving" voice-guided message. Otherwise, voice-guided advice is provided to encourage improving driving behaviors.

$$score_{total} = \left(\sum_{n}^{N} score_{scan,n} \right)/N \times score_v \qquad (1)$$

In previous research, the authors proposed a non-real-time automatic driving behaviors evaluation system [11], which has been used for over 3,000 participants in safe driving courses for licensed drivers conducted by a driving school. Through the safe driving courses, instructors pointed out that participants would found it more intuitive to understand the driving evaluation results when expressed in five grades from A to E, rather than as continuous numerical scores. Similarly, instructors noted that in the context of safe driving courses, debates often arise regarding whether scanning behaviors were performed. They emphasized the importance of not only confirming that scanning behaviors were performed but also communicating when they were insufficient and to what degree they were lacking. Therefore, this study employs a policy of providing voice-guided feedback through a five-level evaluation, rather than continuous scores or a binary good/bad evaluation.

The voice-guided advice consists of as follows: (1) Alert tone to capture attention, (2) Overall performance grade at the evaluation spot, (3) Advice messages for each evaluation items as shown in Table 3. The system provides maximum of two prioritized advisory messages to promote safer driving, based on the driver's failure to perform evaluation items listed in minimum safe-driving behavior checklist at the spot.

4. Experiment and Result

Firstly, to validate the feasibility of measuring visual scanning behaviors using a vision-based driver monitoring sensor, we conducted an indoor experiment with 10 participants (average age: 22.6, $S.D.$ = 0.8). During the experiment, participants were seated 1.2 meters from a wall with markers positioned to correspond with facial orientation angles of 0° (front), 15°, 30°, 45°, and 60° in both left and right directions, taking into account the range of θ_{deg} outlined in Table 1. Participants were instructed to look at each marker from the front position and then return their gaze to the front, continuously repeating this action for each marker twice. The head yaw movements during these tasks were recorded using both a driver monitoring sensor and a gyro sensor (ATR-Promotions, TSND121). For gyro sensor data, facial orientation angles were calculated by time-integrating the angular velocity using the method of previous work [11]. Regarding the driver monitoring sensor, the directly outputted facial orientation angles were utilized. The maximum absolute values of facial orientation angles at the time of looking at each marker were then compared with the true values.
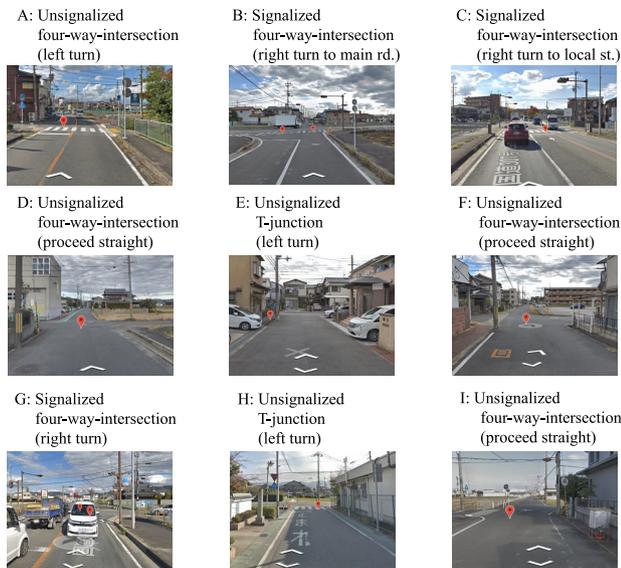
The analysis showed that the mean absolute error (MAE) for the driver monitoring sensor, which directly measures facial orientation angle, was 5.9 ($S.D.$ = 4.2). In contrast, for the gyro sensor, which calculates facial orientation angle by time-integrating angular velocity, MAE was 12.4 ($S.D.$ = 7.5), influenced by the accumulation of errors over time.

Then, we conducted safe driving behaviors evaluations and provided real-time voice-guided safety advice using our proposed system for 27 male (average age: 44.6, $S.D.$ = 12.3) who participated in a safe driving course hosted by Yamashiro Driving School and gave informed consent. In the safe driving course, each participant was instructed to drive a pre-set 5-kilometer route using a driving school car (1500 cc, automatic transmission) equipped with a secondary brake pedal for the professional driving instructor in the passenger seat. The route included nine intersections deemed potentially dangerous spots by the instructor (three

Table 5 Accuracy of proposed system[†].

Evaluation Items	No. of Actions	Evaluation as Good				Evaluation as Bad			
		System	Instructor	precision	recall	System	Instructor	precision	recall
Scanning left before entering spot	189	105 (90)	99	85.7%	90.9%	84 (75)	90	89.3%	83.3%
Scanning right before entering spot	216	120 (115)	132	95.8%	87.1%	96 (79)	84	82.3%	94.0%
Scanning left during turning spot	54	51 (50)	50	98.0%	100%	3 (3)	4	100%	75.0%
Scanning right during turning spot	81	30 (28)	44	93.3%	63.6%	51 (35)	37	68.6%	94.6%
Appropriate speed maintenance	189	146 (127)	132	87.0%	96.2%	43 (38)	57	88.4%	66.7%
Total	729	452 (410)	457	90.7%	89.7%	277 (230)	272	83.0%	84.6%

The figure in parentheses indicates the number of results consistent with the instructor's subjective evaluation.

A: Unsignalized four-way-intersection (left turn)

B: Signalized four-way-intersection (right turn to main rd.)

C: Signalized four-way-intersection (right turn to local st.)

D: Unsignalized four-way-intersection (proceed straight)

E: Unsignalized T-junction (left turn)

F: Unsignalized four-way-intersection (proceed straight)

G: Signalized four-way-intersection (right turn)

H: Unsignalized T-junction (left turn)

I: Unsignalized four-way-intersection (proceed straight)

Source: Google Map

Fig. 3 Selected potentially dangerous spots.

with signals, six without), as indicated in Fig. 3. Table 1 details the minimum safe-driving behaviors at each intersection, typically encompassing two to four evalutaion items per spot.

In this study, we regarded driving behavior as a sequence of driving actions, with each action corresponding to an evaluation item in Table 1. Therefore, one participant's driving behavior over the pre-set 5-kilometer route consists of 27 driving actions. Here, the five-grade evaluation results from A to E comprehensively evaluates the evaluation items listed in Table 1. Thus, the lack of visual scanning behaviors results in a lower grade, however even if visual scanning behaviors are performed, a high vehicle speed will also lead to a lower grade. Therefore, in this study, instead of using grades, we calculate precision and recall for each evaluation item separately to determine the accuracy of evaluation for both visual scanning behaviors and vehicle speed.

Participants' driving actions (729 in total: 27 participants × 27 actions/participant), corresponding to Table 1's evaluation items, were recorded and analyzed by our system, which also provided real-time voice-guided safety advice at potentially dangerous spots. Upon completing the driving, participants were asked to fill out a questionnaire evaluating the system's advisory information and the effectiveness

of the real-time safety advice provided. Additionally, two cameras captured the drivers and traffic conditions for the reference of the instructor's subjective evaluation.

In the subjective evaluation procedure, we requested the driving instructor, who holds a national certification to administer the driving skill tests required for obtaining a new driving license, to subjectively evaluate the participants' driving actions. The evaluation focused on whether the each evaluation item associated with the driving actions was satisfied. For example, regarding the evaluation item "Scan left side-mirror before entering spot", the instructor watches each participant's video and rates it as good if deemed performed correctly, or as risky otherwise. The instructor's subjective evaluation, considered as the definitive evaluation standard in this study, were based solely on these video recordings and not on our system's data. We only showed the driving instructor our experiment's video data and never shared our system's evaluation results nor parameters θ_{deg}, θ_t and θ_v in Table 1. The subjective evaluation is conducted by the same driving instructor.

The system's evaluation for each driving action is conducted using the evaluation item corresponding to the action as listed in Table 1. If all the threshold parameters set (e.g. θ_{deg}, θ_t) for the evaluation item are satisfied, it is judged as good; otherwise, it is deemed as risky. Our system rated 452 actions as good and 277 as risky, detailed in Table 5[†]. Turning to the evaluation result of the driving instructor, he evaluated 457 actions as good and 272 as risky. Of the 452 good actions identified by our system, 410 matched with the instructor's evaluation, yielding a precision ratio of 90.7% and a recall ratio of 89.7%. Similarly, for the 277 risky actions recognized by our system, 230 were in agreement with the instructor, resulting in a precision ratio of 83.0% and a recall ratio of 84.6%.

A closer analysis of accuracy by evaluation item found out that "scanning right during turning spot" (especially scan right-rear side during turning spot at spot B, C and G) had the lowest accuracy. As shown in Fig. 1 and Fig. 4, a vision-based driver monitoring sensor was positioned to the front left of the driver. This placement hinders the sensor's ability

[†]In Table 5, owing to space limitation, multiple evaluation items pertaining to the scanning to the left side before entering spots have been consolidated into a single entry titled 'Scanning left before entering spot'. A similar approach has been applied to the evaluation items for scanning to the right side and scanning during turning spot.

Fig. 4 An example of system evaluation failure for right-rear side scanning during turning spot.

Table 6 Result of Questionnaire.

	Average Score	$S.D.$
Did the driving evaluation enable you to objectively understand your driving behaviors?	4.3	0.72
Did the advice encorage you to drive safer?	4.4	0.64
Did the advice from the proposed system is beneficial for safe driving?	4.5	0.64

to detect facial features like eyes or nose when the driver checks the vehicle's right-rear quadrant, thus challenging the precise determination of facial orientation angles.

The results of the questionnaire using the five-point Likert scale (1:worst–5:best) regarding the proposed system are presented in Table 6. The questionnaire results indicate favorable ratings, with average scores exceeding 4.5 points for questions assessing whether participants could understand their driving, felt motivated to improve upon receiving advice, and found the advice useful for safe driving.

5. Discussion

In the proposed system, potentially dangerous spots are selected, and a checklist of minimum safe-driving behavior for the spot is manually created by driving instructors. However, this manual setup poses operational challenges when instructors are not available, questioning its broader applicability. Consequently, the current implementation of our system is limited to scenarios where instructors are available. To enhance its applicability, this section investigates potential of automated methods for estimating potentially dangerous spots. Interviews with instructors during the creation of the checklist revealed that they focus on several factors when identifying potentially dangerous spots: (1) intersection shape, (2) width of roads connecting to the intersection, (3) presence of traffic signals, (4) existence of stop sign regulations and (5) visibility distance at the intersection. Thus, as shown in Table 1 and Fig. 3, intersections sharing these characteristics have the same checklist for minimum safe-driving behavior (e.g., spots D and F). Therefore, if factors (1) through (5) can be estimated, it becomes feasible to auto-generate a checklist of minimum safe-driving behavior for the spot by cross-referencing Table 1.

Among the above-mentioned factors, (1) to (4) can be automatically detected using road network databases that encompass the entire country (e.g. MapFan®DB, GeoTech-

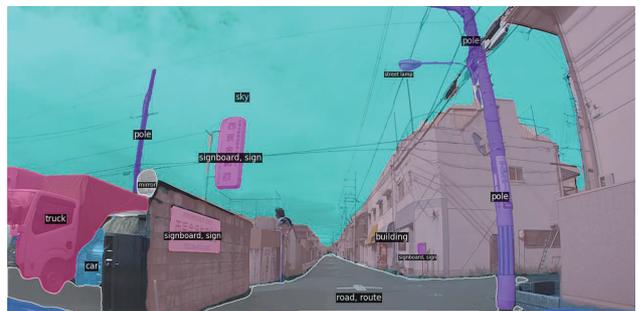Input image

Segmentation result

Fig. 5 An example of semantic segmentation.

nologies, Inc.). These databases contain detailed intersection data (such as latitude, longitude, traffic regulations, and the number of roads connecting to the intersection) as well as information about the roads leading to each intersection (including width, number of lanes, and traffic regulations). For (5), recent rapid advancements in semantic segmentation [12], a technology that classifies each pixel in an image based on semantic meaning, enables automatically detecting road sections and buildings around intersections. Figure 5 shows an example of applying OneFormer [13], a type of semantic segmentation, to an image of intersection, successfully detecting objects such as buildings around the intersection that affect visibility distance. As demonstrated, combining road network databases with semantic segmentation holds the possibility for the future automatic estimation of potentially dangerous spots.

6. Conclusion

Most traffic accidents are due to driver error. Thus, addressing drivers' errors is as crucial as improving vehicles and road infrastructure. This paper introduces a real-time safe driving advisory system utilizing a vision-based driver monitoring sensor to detect drivers' visual scanning behaviors. The system evaluates safety driving behaviors using criteria based on professional driving instructors' expertise, providing voice-guided safety driving advice to promote safer driving practices.

In real traffic environment experiment involving 27 drivers, our system aligned with driving instructors' safety behaviors evaluations with over 80% accuracy. Currently, we are piloting our system in a driving school to aid profes-

sional drivers' retraining. Our future goal is to determine which types of advice most effectively encourage safe driving during driver retraining.

Acknowledgments

References

[1] M. Hasenjäger, M. Heckmann, and H. Wersing, "A survey of personalization for advanced driver assistance systems," IEEE Trans. Intell. Veh., vol.5, no.2, pp.335–344, 2020.

[2] J. Wang, W. Chai, A. Venkatachalapathy, K.L. Tan, A. Haghighat, S. Velipasalar, Y. Adu-Gyamfi, and A. Sharma, "A survey on driver behavior analysis from in-vehicle cameras," IEEE Trans. Intell. Transp. Syst., vol.23, no.8, pp.10186–10209, 2022.

[3] M.H. Baccour, F. Driewer, T. Schäck, and E. Kasneci, "Comparative analysis of vehicle-based and driver-based features for driver drowsiness monitoring by support vector machines," IEEE Trans. Intell. Transp. Syst., vol.23, no.12, pp.23164–23178, 2022.

[4] A. Kashevnik, R. Shchedrin, C. Kaiser, and A. Stocker, "Driver distraction detection methods: A literature review and framework," IEEE Access, vol.9, pp.60063–60076, 2021.

[5] Y. Qu, H. Hu, J. Liu, Z. Zhang, Y. Li, and X. Ge, "Driver state monitoring technology for conditionally automated vehicles: Review and future prospects," IEEE Trans. Instrum. Meas., vol.72, pp.1–20, 2023.

[6] A. Misra, S. Samuel, S. Cao, and K. Shariatmadari, "Detection of driver cognitive distraction using machine learning methods," IEEE Access, vol.11, pp.18000–18012, 2023.

[7] T. Hirayama, K. Mase, C. Miyajima, and K. Takeda, "Classification of driver's neutral and cognitive distraction states based on peripheral vehicle behavior in driver's gaze transition," IEEE Trans. Intell. Veh., vol.1, no.2, pp.148–157, 2016.

[8] T. Tanaka, K. Fujikake, T. Yonekawa, M. Yamagishi, M. Inagami, F. Kinoshita, H. Aoki, and H. Kanamori, "Study on driver agent based on analysis of driving instruction data—Driver agent for encouraging safe driving behavior (1)—," IEICE Trans. Inf. & Syst., vol.E101-D, no.5, pp.1401–1409, 2018.

[9] T. Tanaka, K. Fujikake, T. Yonekawa, M. Inagami, F. Kinoshita, H. Aoki, and H. Kanamori, "Effect of difference in form of driving support agent to driver's acceptability—Driver agent for encouraging safe driving behavior (2)—," J. Transp. Technol., vol.8, no.3, pp.194–208, 2018.

[10] K. Chen, T. Yamaguchi, H. Okuda, T. Suzuki, and X. Guo, "Realization and evaluation of an instructor-like assistance system for collision avoidance," IEEE Trans. Intell. Transp. Syst., vol.22, no.5, pp.2751–2760, 2021.

[11] M. Tada, H. Noma, A. Utsumi, M. Segawa, M. Okada, and K. Renge, "Elderly driver retraining using automatic evaluation system of safe driving skill," IET Intell. Transp. Syst., vol.8, no.3, pp.266–272, 2014.

[12] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, "Review the state-of-the-art technologies of semantic segmentation based on deep learning," Neurocomputing, vol.493, pp.626–646, 2022.

[13] J. Jain, J. Li, M. Chiu, A. Hassani, N. Orlov, and H. Shi, "One-Former: One transformer to rule universal image segmentation," arXiv:2211.06220, https://arxiv.org/abs/2211.06220, 2022.