IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 31, 2021, accepted June 16, 2021, date of publication June 29, 2021, date of current version July 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3093337

A Machine Learning-Based Approach to Analyze Information Used for Steering Control

YUKI OKAFUJI[1], (Member, IEEE), TOSHIHITO SUGIURA[1], RYUSEI OSUGI[1], CHENKAI ZHANG[1], AND TAKAHIRO WADA[2], (Member, IEEE)

[1]College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan
[2]Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan

Corresponding author: Yuki Okafuji (yokafuji@fc.ritsumei.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 20K14996, and in part by the Suzuki Foundation.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Ritsumeikan University under Reference Number BKC-JinI-2019-021.

ABSTRACT Understanding the relationship between driving behavior and visual information is important for holistic understanding of driving behavior. However, the analysis of the cognitive behavior for steering/throttle control has been only conducted under a special simulator environment. Therefore, in this study, we aimed to develop a convolutional neural network (CNN) with human physical characteristics to analyze the driver's cognitive behavior and to validate that the machine learning methods can be an analytical method for understanding driver behavior. We obtained the driving data in a simulator experiment to train the proposed CNN model. The region where the visual field influences drivers' steering behavior was analyzed using the results of the feature maps generated by the trained CNN model and the driver's gaze behavior. The results indicate that the driver performs steering control using the information within 20 degrees from the gaze point. This shows that the results obtained from our proposed method can reproduce the same results as previous findings. We also validated that the results are not uniquely obtained depending on the proposed model and environment but are also influenced by the driving behavior such as the gaze point and the steering control. We analyzed the dataset generated by the mathematical control model, called the driver model, which performs different behaviors from the driver. The analysis results generated by the driver model were different from the results of the human data. Therefore, the results generated by the machine learning-based analysis are influenced by the driving behavior. Consequently, these results imply that machine learning methods have the potential to become analytical methods for understanding driver behavior.

INDEX TERMS Cognitive behavior, convolutional neural network, human vision, steering behavior.

I. INTRODUCTION

Humans use their visual system to operate a car skillfully through the processes of perception, cognition, decision-making, and operation [1]. Understanding these processes of driving behavior is important not only from a scientific point of view, such as understanding human behavior, but also from the engineering point of view, such as safe vehicle design. To understand holistic driving behavior, it is not enough to understand only one aspect of this behavior; examining all behavioral phases is important. In recent decades, driver behavior has been analyzed mainly through psychological experiments to verify each aspect of drivers' characteristics.

For example, while researching perception, many methods have elucidated the information perceived by the driver by measuring the driver's gaze points [2]. The gaze points are relatively easy to estimate using measuring devices such as eye trackers, and lately, gaze points can be measured accurately using area of interest (AoI) [3] as the method of analysis, and optokinetic nystagmus (OKN) [4]. The drivers' gaze behavior has been analyzed not only on relatively simple roads, but also in many situations such as intersections where driving is more difficult [5]. Regarding research on decision-making, a mathematical model called a driver model is often used [6]. Driver models with various perceptual information and structures are compared to the data obtained from the experiment to elucidate the decision-making process for the operation. Several models such as the preview control model [7] and the two-point steering control model [8] have been proposed.

The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li.

As more meta-decision-making research, the task of classifying driving behaviors such as following and lane changing has also been studied vigorously [9]. These are also applied as control methods for automated driving that considers human characteristics [10]. Research on operation has been validated for a very long time. By analyzing easily observable data such as the driver's steering and throttle operation, the characteristics of individual drivers can be analyzed. Various operating behaviors have been verified, such as differences in steering torque performance [11], vehicle running path [12], and differences in the operation input between general and racing drivers [13].

While each behavioral phase has been examined as described above, cognitive behavior has also been analyzed. Cognitive behavior has various aspects, of which is attention often examined. Attention as an aspect of driving is defined as the process of bringing an object to awareness [14]. In terms of attention studies, based on the notion of a useful field of view (UFOV), the regions of central and peripheral vision are often analyzed depending on the difference in reaction time [14]. Extending the concept of UFOV, there are many verifications of attention during driving. For instance, the attention to age difference [15], in specific environments such as intersections [16], and the object recognition rate rather than reaction time [17] have been analyzed in previous studies.

Meanwhile, studies on cognitive behavior for decision-making, and steering/throttle control aim to clarify the kind of information and the region the visual field drivers use for their operations. For the kind of information used for control, road edges and retinal flow are used for steering control [18], [19]. For the region in the visual field used for the control, two regions, the far and near regions from the vehicle position as the two-point steering control model [8], predominantly affect steering behavior [20]–[22]. In these studies, cognitive behavior has been commonly analyzed by intentionally hiding information about a specific region, such as the far and near regions, in simulator experiments. However, this analytical method has some limitations like the discussion of the generality from driving behavior obtained under unusual conditions, although it can provide an understanding of useful driver behavior. This is different from other analytical methods of analyzing perception, decision-making, and operational behaviors. There are no general analytical methods to elucidate the cognitive behavior of steering and throttle control. Therefore, analytical methods beyond the framework of conventional psychological experiments are required.

In recent years, the concept of attention has been widely used in machine learning techniques for classification and regression tasks in image processing. The attention mechanism visualizes the image region affecting the tasks. This visualized region is called a feature map. The feature map in the classification task is related to the output location of the recognized object in the image [23]. Moreover, the regression task is often used in the field of automated driving. The feature maps visualized in the input image strongly influence the computation of output in a steering model with imitation learning of human driving behavior [24]–[26]. This visualized image allows debugging to clarify the decisions of automatic systems. The attention mechanism of machine learning is not only used for inference in black-boxed deep learning models, but also beginning to be applied to the analysis of human behavior. For instance, the results of attention as machine learning have been applied to an analytical method to analyze attention as the driver's cognitive behavior [27], [28]. This means that there is a possibility of applying this to the analysis of cognitive behaviors related to driving behavior itself, such as steering and throttle control.

This study aims to develop an analysis method for cognitive behavior with a focus on steering control based on machine learning techniques. In particular, the main focus of analyzing cognitive behavior is to clarify the region of the visual field that drivers use for their operations. However, it is unclear that the results obtained from the machine learning method are really interpreted as features of the driver, as machine learning is often considered to have black-box features. Therefore, in this study, we follow the steps below to validate whether the results obtained from the machine learning method can reflect driver behavior. This allows us to validate that machine learning methods have the potential to become analytical methods for understanding driver behavior.

1) Development of a convolutional neural network (CNN) model with human physical characteristics
2) Analysis of driver behavior obtained in the simulator study using the proposed model
3) Assessing the validity of results as cognitive behavior using driver models

The contributions of this study are as follows:

- We propose a machine learning method including human physical characteristics to analyze the region of the visual field where drivers use for their steering operation. This method does not require any special simulator environment such as intentionally hiding certain areas, which has been conducted in previous studies [8], [19], [20], [22].
- We verify that the proposed model can reproduce driving behavior and have the same results in previous studies. In addition, we show that the results generated by our proposed model can reflect the driving behavior of the measured participants, with an analysis using the driver model. These analyses imply that machine learning methods have the potential to become analytical methods for understanding driver behavior.

In this paper, in Section II, the development of the proposed CNN with human physical characteristics is described. Section III provides an overview of the dataset generated by the simulator experiments. Section IV describes the analysis results of the proposed model. In Section V, we validate the results obtained from machine learning techniques, which can be interpreted as the cognitive behavior of the driver. Section VI and VII provide the limitations and conclusions of this study, respectively.

The preliminary research for this study was presented at a conference and published in its proceedings [29]. We reported the comparison of the fundamental theoretical frameworks suitable for analyzing driver data of only one participant in this preliminary study. The current paper has been refined accordingly, and provides a detailed analysis in Section IV and describes validation results in Section V.

II. PROPOSED ANALYSIS MODEL

In this study, we use PilotNet as the machine learning method for analyzing the driver's cognitive behavior [24], [25]. PilotNet is one of the CNN models, and a model that calculates the steering output from the image input in front of the vehicle. We consider PilotNet to be suitable for this study because it works very well as an automated driving controller that imitates driver operations, and has a structure to visualize feature maps that influence the operations. However, there are several inadequacies in using PilotNet as an analysis method for human cognitive behavior. For instance, humans have various structural delays from perception to the operation process. As PilotNet was developed for automatic driving, there is no human-like structural delay in the computation of inputs and outputs. This structure is significantly different from the physical characteristics of humans, and is, therefore, insufficient to analyze a driver's cognitive behavior. Hence, in this study, we propose a CNN model for human analysis that considers the physical characteristics of humans.

As explained in the introduction, various driver models have been proposed to analyze the characteristics of drivers (e.g., [30]). The driver model considers human physical characteristics such as time delay and feedback systems. Therefore, in this study, we propose a CNN-based driver model (FB-Delay-PilotNet) that introduces these time delays and feedback system to PilotNet to represent the cognitive process of the driver (Fig. 1). The proposed model includes two types of human-like time delays: the processing time delay, in which sensory information is transmitted to the human brain; and the neuromuscular dynamics, when the brain output uses nerves to move muscles. Accordingly, we implement the dead time system as a processing time delay and a first-order control system simulating neuromuscular dynamics into FB-Delay-PilotNet. The coefficients T and τ in the dead time and first-order control systems are 40 ms and 100 ms [30], respectively. For the feedback system, the steering angle, one step before the current state, is added to the output layers. Such a feedback system can be interpreted as the process of calculating the adjustment output for the current steering angle state.

FB-Delay-PilotNet is composed of 11 layers, as shown in Fig. 1: a normalization layer of the input image, five convolutional layers, a dropout layer, three fully connected layers, and an output layer. The size of the input image is set to 295 × 800 pixels, depending on the dataset. The difference between the original PilotNet and our proposed model, except for the structure representing the human physical characteristics, is the size of each layer associated with differences in the size of the input images. The details of the layers in FB-Delay-PilotNet are shown in Table 1. The loss function $L(x)$ (MSE: mean square error) to simulate driver steering behavior is represented by (1):

TABLE 1. Details of layers in FB-Delay-PilotNet.

Layers	Output Size	(Channel, Kernel, Stride)	Activation Function
Convolution 1	(146, 398)	(24, 5, 2)	ELU
Convolution 2	(71, 197)	(36, 5, 2)	ELU
Convolution 3	(34, 97)	(48, 5, 2)	ELU
Convolution 4	(16, 48)	(64, 5, 2)	ELU
Convolution 5	(7, 23)	(64, 5, 2)	-
Layers	**(Input Size, Output Size)**		**Activation Function**
Fully Connected 1	(103040, 1000)		ELU
Fully Connected 2	(1000, 250)		ELU
Fully Connected 3	(250, 50)		-
Output	(51, 1)		-

$$L(x) = \frac{1}{M} \sum_{m=1}^{M} \left(\hat{\delta}_c^m(x) - \delta_c^m(x) \right)^2, \tag{1}$$

where x represents the input image, $\delta_c(x)$ denotes the value inverted from the observed steering angle through the first-order control system, and $\hat{\delta}_c(x)$ indicates the estimated value calculated from FB-Delay-PilotNet. M is the batch size (128), and m means the element of each value included in the mini-batch. The optimized algorithm was set to Adam (learning rate: 0.0001).

To improve the interpretability of PilotNet after training the dataset, VisualBackProp (VBP) has been proposed in a previous study [25]. The VBP can visualize the regions in the input image that have a large effect on the output using weights in the convolutional layers. The overview of the VBP algorithm is that the average and deconvolutional feature maps in each convolutional layer are multiplied until the result becomes the same size as the input image, as shown in Fig. 1. In this study, we also apply the VBP algorithm to the proposed FB-Delay-PilotNet. Then, we can analyze the specific regions in the input that are associated with the steering performance of the human driver.

The FB-Delay-PilotNet proposed in this study has been compared to the original PilotNet in our preliminary study [29]. The result shows that the proposed model can simulate the driver's steering behavior with higher accuracy than baseline PilotNet. Thus, we apply the FB-Delay-PilotNet for a human cognitive analysis in this study.

III. DATASETS OF HUMAN DRIVER

In this section, we analyze the cognitive behavior responsible for the driver's steering operation, using the proposed model. This study aims to establish an analysis method for cognitive behavior using machine learning techniques. Thus, in the first step, we conducted a simulator experiment, to ensure easy control of the environment, and collected data. Then,

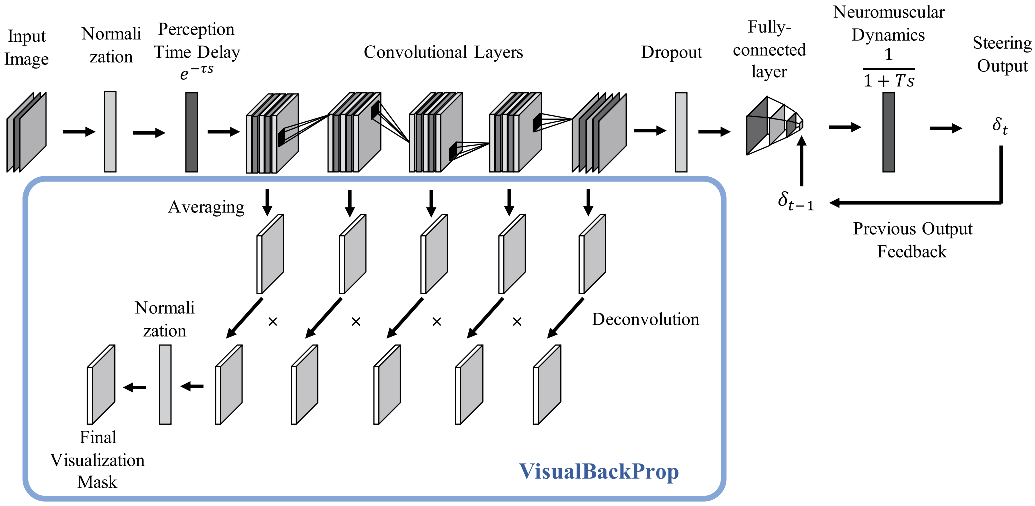


FIGURE 1. Structure of proposed FB-Delay-PilotNet and visualization method.

we show that the results obtained from our proposed method, trained by collected data, can reproduce the same results as previous findings, which imply that the machine learning method could be an analysis method for driving behavior.

A. ETHICS

This study was approved by the Research Ethics Committee of Ritsumeikan University (Reference number: BKC-JinI-2019-021). It complied with all the guidelines in the Declaration of Helsinki.

B. PARTICIPANT

The sample consisted of three University students (3 males, 20, 22, and 21 yrs) who took part in this study. All participants had normal vision and had held a driving license for one year. The participants signed informed consents, which allowed for the use of the collected data for scientific purposes and publication. The participants received 1,000 JPY per hour for their participation.

C. APPARATUS

The driving simulator used in our experiment is shown in Fig. 2. The virtual environments were generated using the Vizard 5.0 (WorldViz) software on a PC (Intel Core i7-8700 CPU), and projected (BenQ TH671ST). The projection size was 2.435 m × 1.36 m; the participant sat 1.6 m from the screen; and their eye height was 1.5 m; thus, the field of view was 74.5 × 46.1 deg. The steering wheel was controlled using a Logitech G29 wheel (Logitech). The simulation was run at 40 Hz. To collect the gaze data, the participant wore Tobii Pro Glasses 2 (Tobii Technology K. K.) sampled at 50 Hz.

D. EXPERIMENTAL DESIGN

Real-world driving lanes inherently have various straight and curved sections that induce natural driving. In this study, 32 road shapes were used; these were extracted from a real-world map and depicted as 3 m wide in the driving simulator (e.g., Fig. 3). These 32 courses were duplicated and inverted (a total of 64 running courses) to balance the number of left and right curves. The velocity of the vehicle was constant (45 or 90 km/h); thus, only the steering could be manipulated by the participant in this study. Within the course, 34 lanes had large curvatures and were used for the 45 km/h condition, and the remaining 30 had small curvatures and were used for the 90 km/h condition. The average running times for the 45 and 90 km/h conditions were approximately 215 and 150 s, respectively. As the seat position influences driving behavior [3], [32], [33], the simulated seat position was set at the center of the vehicle. The participant was instructed to drive on the center of the road with smooth and accurate steering. Before the experiment, the participants practiced for approximately 7 minutes on the training course to get used to the driving simulator environment. The experiment was conducted in two sessions over two days to reduce the participants' fatigue.

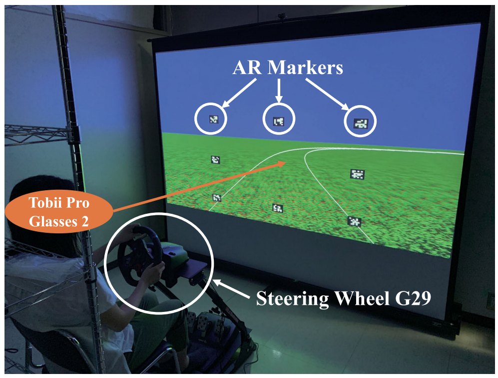


FIGURE 2. Driving simulator. Eight augmented reality (AR) markers were set on the screen to calculate the driver's gaze.

E. GENERATE DATASET

The vehicle information (position, orientation, velocity, and angular velocity), input steering angle, and gaze points on the

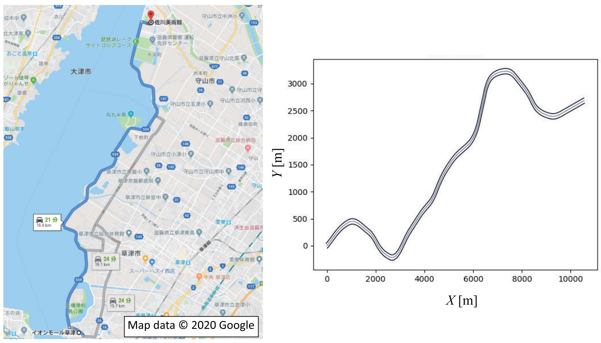


FIGURE 3. Example of a driving course. The left figure shows a map retrieved from Google Maps, and the right figure is a road simulating the real-world maps.

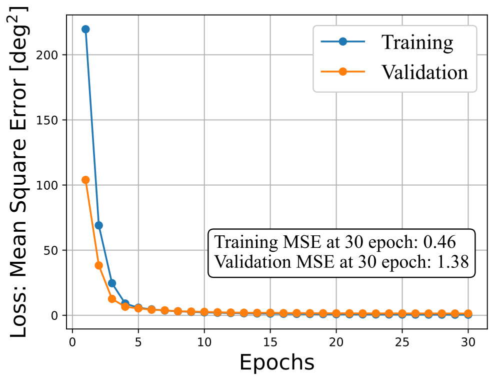


FIGURE 4. Average value of the loss function in training/validation data for Participant 1's dataset.

screen were obtained in the experiment. To calculate the gaze points on the screen, we used Tobii Pro Glasses 2 and AR markers on the screen (Fig. 2). The calibration method and precision of Tobii Pro Glasses 2 are described in the official document [34]. All measurement data were resampled to 60 Hz, and the measurement time was synchronized. The simulation screen images that excluded information unrelated to steering behavior (AR markers and sky information) were captured and used for training the proposed FB-Delay-PilotNet. The size of the training image was 295 × 800 pixels. The FB-Delay-PilotNet was trained by using a combination of the input screen images and steering angles as outputs.

IV. EXPERIMENTS AND RESULTS

A. ACCURACY OF LEARNING MODELS

In this study, the proposed FB-delay PilotNet is performed for training and validation using data from 58 out of 64 trials conducted in the experiment. The remaining six trials, with medium- and high-speed conditions, are used as test data. Additionally, when training the proposed model, 90% of the 58 trials of training data are used for training, and the remaining 10% are used for validation.

Fig. 4 shows the average values of training and validation losses through 30 epochs for Participant 1's dataset. Moreover, to show the accuracy to represent steering performance in detail, Fig. 5 indicates the plots of the steering angle for one course in each velocity within the test data of Participant 1. Table 2 shows the MSE results for all participants at different velocities. From these results, it can be confirmed that after training, the FB-Delay-PilotNet accurately represented the driver's steering performance.

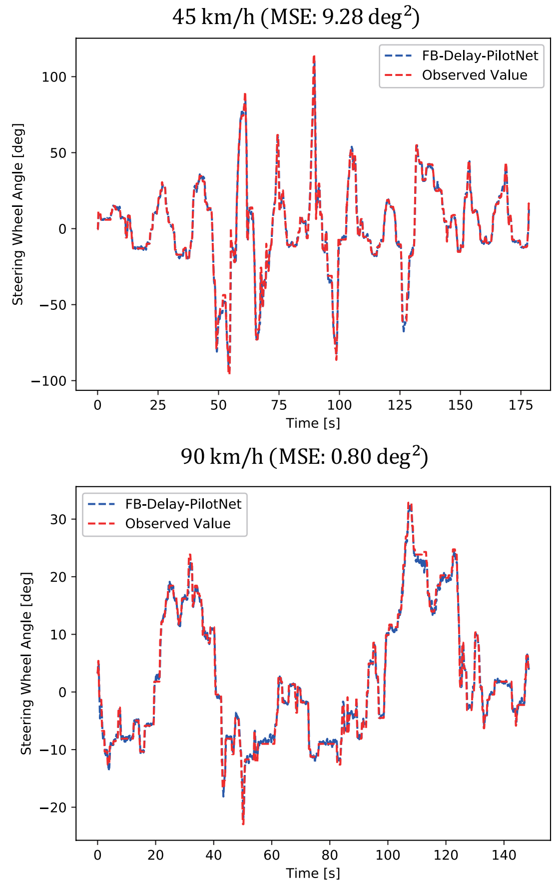


FIGURE 5. The results of the observed steering angle (red line) and steering angle produced by the FB-Delay-PilotNet (blue line) in test data of Participant 1.

TABLE 2. MSE results for all participants.

Participants	MSE deg^2 in 45 km/h	MSE deg^2 in 90km/h
1	9.28	0.80
2	7.60	1.35
3	14.61	0.33

B. VISUALIZING IMPORTANCE BY VISUALBACKPROP

After training the FB-Delay-PilotNet, we visualize an important area in the input images, which influences the steering performance, using the VBP algorithm. Fig. 6 shows an example of the result of the feature map using the VBP in test data of Participant 1. Regions with higher weights, which are associated with steering behavior, are shown on the road edges and the textures in the area close to the vehicle.

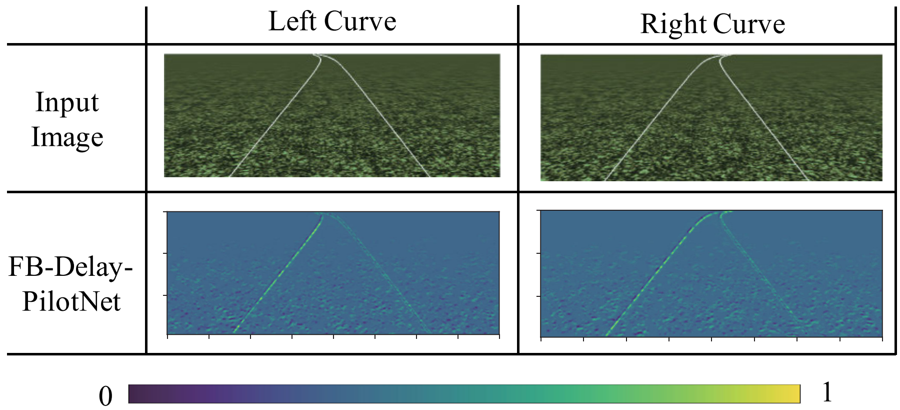


FIGURE 6. Examples of feature map by VisualBackProp (VBP) in test data of Participant 1. Pixels with higher values in the image have a strong influence on the steering performance.

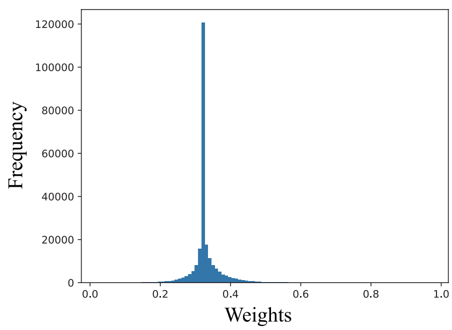


FIGURE 7. Example of a histogram of weight in the feature map for one frame in 45 km/h.

In the VBP analysis, the distribution of weights in the feature maps among all frames differs according to the road and vehicle state. Accordingly, to evaluate the weight importance through all driving data with the same criteria, we first depict the properties of the frequency of occurrence of weights in each frame. Fig. 7 shows an example of a histogram of weight in the feature map for one frame in the medium-speed condition. This histogram has a shape similar to the Gaussian distribution. Therefore, the threshold processing represented in (2) is performed to equalize the pixels with important weights between all the frames.

$$W(h,w) = \begin{cases} 0, & (W(h,x) < \mu + n\sigma) \\ W(h,x) - (\mu + n\sigma), & (otherwise), \end{cases} \tag{2}$$

where the average of the weights obtained in each frame is μ, the standard deviation of the weights denote σ, the weight of (h, w) pixels in the feature maps by the VBP is represented as $W(h, w)$, and n indicates a threshold parameter.

Fig. 8 shows an example of the feature map with threshold process when $n = 0, 1, 2, 3$ in (2). Areas in the image except for with weight 0 strongly influence the steering performance. If we assume the weight histogram shown in Fig. 7 as a Gaussian distribution, $n = 0, 1, 2, 3$ means extracting the top 50.0 %, 15.9 %, 2.3 %, 0.1 % of important information in the input image, respectively. From the results in Fig. 8, when n increases, only the importance of information regarding the road edges are included in the feature maps.

C. GAZE DISTRIBUTION

Fig. 9 shows the distribution of gaze points of Participant 1 on the screen in both speed conditions, which is the same image size as the feature maps. This result shows that the distribution of gaze points in the medium-speed condition is wider in the horizontal direction than that in the high-speed condition. Meanwhile, in the high-speed condition, the gaze points are concentrated towards more distant areas. Comparing with Fig. 6, it can be confirmed that the gaze points are distributed near the center of the road, and the only difference between the speed conditions is in the gaze distance.

In our previous study [35], we show a detailed analysis of the driver's gaze behavior using the driving data measured in this experiment. The results show that drivers' gaze was influenced by the vehicle's state. When their vehicle tracking control was stable, drivers tended to fixate towards the center of the road at farther areas. Whereas, when lane-keeping control was strongly required, drivers tended to fixate at points closer to the road edges at nearer areas.

D. RELATIONSHIP BETWEEN FIXATION AND FEATURE MAPS

This study aims to investigate drivers' cognitive behavior. This purpose can be verified by comparing the feature maps shown in Fig. 8, which indicates which areas of the input image are important for steering performance, and where the participant gazes at the feature maps. First, instead of (2), we recalculate the important areas in the original feature maps that affect the steering performance as follows:

$$W(h,w) = \begin{cases} 0, & (W(h,x) < \mu + n\sigma) \\ 1, & (otherwise). \end{cases} \tag{3}$$

Second, we superimpose the gaze point on the recalculated feature maps, and calculated the visual angle between the

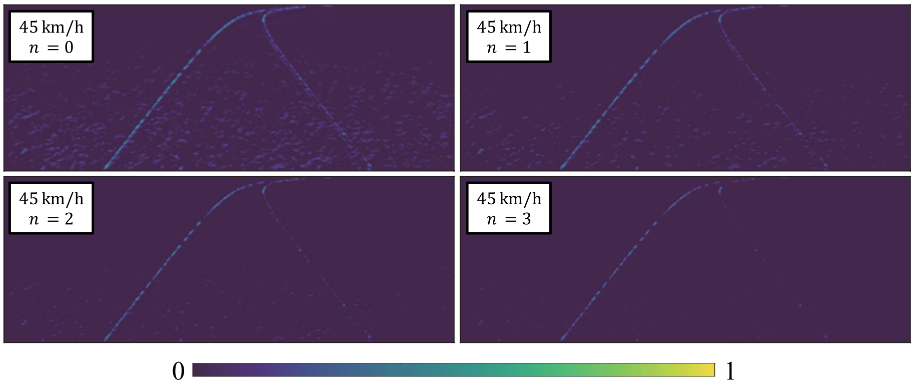


FIGURE 8. Example of feature maps by the VBP with threshold process in 45 km/h.

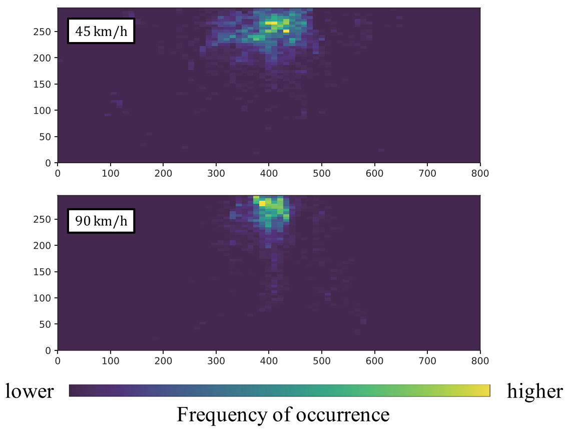


FIGURE 9. Gaze distribution of Participant 1.

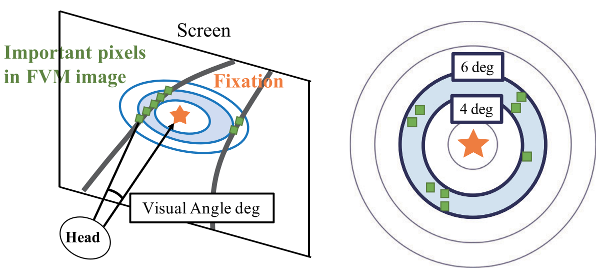


FIGURE 10. The method to calculate the visual angle between the gaze point and each important pixel on the screen. The figure to the right shows the ratio of important pixels (green) to the area (light blue).

gaze point and all pixels with high importance (left figure in Fig. 10). Additionally, when the visual angles for all important pixels are divided by every two degrees, we calculate the weight ratio, which is the number of important pixels (green pixels in Fig. 10) to the area of the divided regions (light blue area in Fig. 10). Finally, the weight ratio in each divided region is calculated through all frames in one test course, and the average and standard deviation of the weight ratio for each visual angle is calculated accordingly.

Fig. 11 shows the results of the weight rate for all participants for each test course in both 45 and 90 km/h. The results except 90km/h for Participant 2 show that the important weights of the feature map exist around 0-20 degrees from the gaze point, regardless of the difference in speed. Moreover, the weight ratio decreases from more than 20 degrees of the visual angle. Furthermore, as the threshold of n increases, the overall ratio decreases, especially at the visual angle of about 10-20 degrees. In the 90 km/h results for Participant 2, the weight rates below 10 degrees are reduced for $n = 1$ and 2. However, for $n = 3$, we can see that the weight rates above 10 degrees are especially reduced. In other words, the trend with increasing n is the same as the other results.

The results can be discussed in relation to the characteristics of human vision. Human vision can be broadly divided into the fovea, central vision, and peripheral vision [31]. The range of the fovea is often determined to be within two degrees from the gaze point, while the range of central and peripheral vision varies depending on the task. For example, in low-dimensional tasks such as measuring reaction time, the central vision is often considered to be within 20-30 degrees from the gaze point [31]. Therefore, the definition of the central vision in this study is also considered 20-30 degrees. Fig. 11 shows that for $n = 1$, the important information is concentrated in the region of the central visual field, and for $n = 3$, the information close to the fovea region becomes more important. Therefore, the proposed model can be considered to show the performance of eliciting characteristics close to human vision. Thus, this result implies that the machine learning method could be an analysis method for driving behavior.

Meanwhile, the results of this study do not show any difference between speed conditions. This can be attributed to the simple simulator environment. In this study, the simulator environment had only the road edges and texture to generate optical flow. In contrast, in a typical driving environment, where there are many cars and pedestrians, more speed-dependent driving behavior is induced [38]. Therefore, the cognitive behavior of the driver depending on speed is not observed in this study.

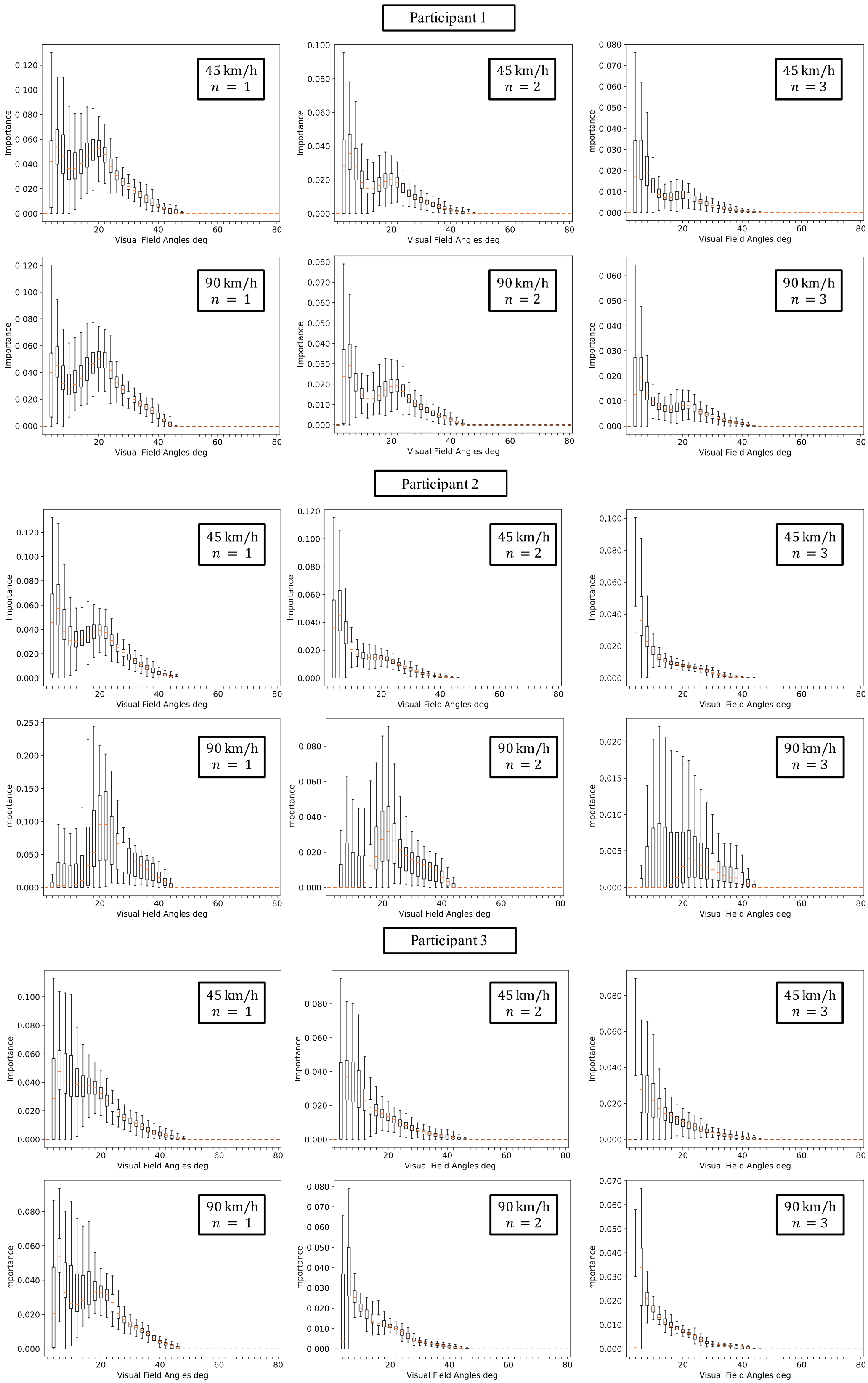


FIGURE 11. Important weight ratio of each information in the visual field according to velocity and threshold value.

V. VALIDATE VISUALIZATION PERFORMANCE BY DRIVER MODEL

In the previous section, an analysis method of the driver's cognitive behavior is proposed by comparing the driver's gaze point and the feature maps generated by the trained FB-PilotNet. This analysis method assumes that the results of the feature map are influenced by the driver's gaze point. However, we cannot guarantee that the feature maps are dependent on driving behaviors, such as the gaze point and driving style. Regardless of the driving behavior, the feature maps may always show the same result. Hence, it could be an incorrect method of analysis of the driver's cognitive behavior.

Therefore, in order to validate whether our proposed model could reflect the driving behavior of the measured participants, we generate a new dataset using a mathematical model called the driver model. The FB-Delay-PilotNet is retrained to compare the new results and the results described in the previous section. The driver model refers to a point different from the participants' gaze point and performs a different control style depending on the reference point. Then, the FB-Delay-PilotNet trained by the new dataset outputs the new results of the feature maps. If the results of the feature maps between the driver model and the participants are the same, it means that the results presented in the previous section coincidentally correspond to human characteristics. Whereas, if the results between the driver model and the participants are different, the results of the FB-Delay-PilotNet can be considered to reflect the participants' driving behavior. In other words, the results of the previous section can be interpreted as one of the results of human cognitive behavior. Thus, we can validate that the machine learning method could indeed be a method of analysis for understanding driving behavior.

A. DRIVER MODEL

The two-point visual control model [8] is a famous driver model that can reproduce human-like steering behavior. This model refers to the far and near regions and uses visual angles between the vehicle direction and the reference points in the far and near regions for steering control. In this study, we use a simplified control method of the two-point model with only one reference point for dataset generation, as shown in Fig. 12. A point on the center of the road, at a constant distance of 8.3 m from the vehicle, was used as the reference point. The steering controller is represented as

$$\dot{\delta} = K_p\dot{\phi} + K_I\phi. \tag{4}$$

where δ represents the steering angle, ϕ denotes the visual angle between the vehicle direction and the reference point, and K_p, K_I are the control gains.

We use this model to generate a new dataset by running the vehicle automatically in the same simulator environment (course and velocity) as described in Section III.

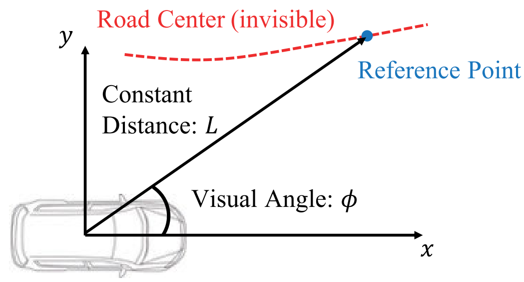


FIGURE 12. Control method referring to only one point. A point on the center of the road, 8.3 m away from the vehicle, was used as the reference point.

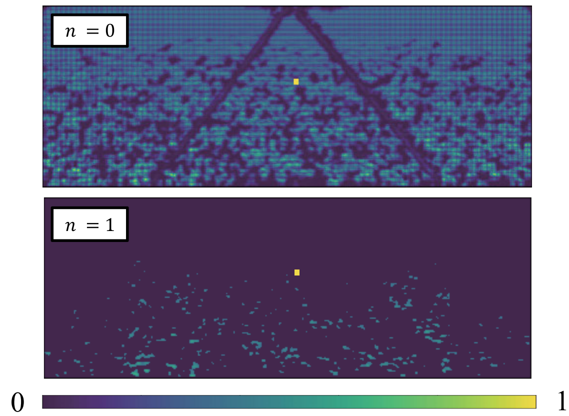


FIGURE 13. Example of feature maps with a threshold process generated by driver model. Yellow points depict the position of the reference point.

B. VISUALIZING IMPORTANCE

We retrain the FB-Delay-PilotNet on the new dataset and output the feature maps of the input image by the VBP algorithm on the test course. Fig. 13 shows an example of the feature maps with $n = 0, 1$ in (2) in medium-speed condition. Comparing Fig. 13 with Fig. 8 generated by the driver data, we confirm that the trends of the feature maps are different. In Fig. 8, the important weights are clustered on the road edges, especially in the regions far from the vehicle. However, in Fig. 13, the important weights are clustered on the texture in the region close to the vehicle. The reason why the important weights are dispersed from the reference point is that a point on the invisible center of the road was set as the reference point of the control model. Therefore, a clear reference point cannot be estimated. From these results, we can confirm that the important weights are clustered in the region around the reference point in Fig. 13, as well as, and that the important weights are also clustered around the gaze point in Fig. 8.

C. RELATIONSHIP BETWEEN REFERENCE POINT AND FEATURE MAPS

Finally, by superimposing the results of the reference points and the feature maps, as in Section IV-D, the results of the weight ratio of important weights based on the reference

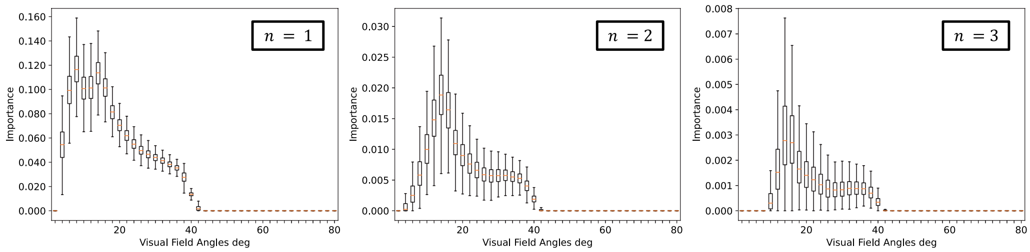


FIGURE 14. Important weight ratio based on the reference point according to the threshold value.

points in the medium-speed condition are shown in Fig. 14. Comparing Fig. 14 with Fig. 11 generated by the driver data reveals that the trends of the weight ratio are also different because the visualization performance is different. In particular, as the threshold n increases, the difference increases. In the driver data, as the threshold value increases, the important weight is concentrated in the region close to the gaze point. However, in the driver model data, the ratio of important weights decreases in the regions closer to the reference point. Thus, these results show a difference between human and driver model data.

When the proposed method is trained on a dataset generated using a driver model, the results are completely different from those shown in Section IV. This is because we use a driver model that is different from the driver's gaze point and steering behavior. This means that the results obtained from the FB-Delay-PilotNet are influenced by driving behaviors, such as gaze point and steering controller. Therefore, the results described in Section IV are not uniquely obtained depending on the proposed model or environment but are influenced by the characteristics of the participants. This implies that the machine learning method could be a method of analysis for understanding driving behavior.

VI. LIMITATION

We consider that this study has several limitations.

- First, we do not confirm whether similar results can be obtained in more complex environments. In this study, the dataset is generated in a simulator environment with only road edges and texture. In Section IV, we explain that there is no difference in speed conditions because the environment is too simple. Human driving behavior is dependent on the surrounding environment [38], this study did not reproduce those driving characteristics in this regard. Thus, we need to examine whether the proposed method can be used as a general-purpose analysis method by validating it in a more complex environment.
- Second, in Section IV, the proposed model is trained by only three participants; and a large difference in the results could be observed between participants. Therefore, the driver model is used to verify whether differences in driving behavior would cause differences in the results of the model as cognitive behavior. In future studies, as individual driving behavior differs greatly in complex environments, individual differences in cognitive behavior need to be verified in complex environments.
- Third, the proposed model does not completely reflect human physical characteristics. For instance, the FB-Delay-PilotNet proposed in this study used only RGB images as input information. However, humans use input information such as optical flow for steering control [19], [36], [37]. Moreover, some studies consider the feedback torque from the steering wheel to accurately introduce the physical structure of the driver [30]. In the future, it will be necessary to consider the extent to which these human characteristics need to be taken into account in the analysis model.
- Finally, it is necessary to evaluate to the extent to which the results of this study reflect human cognitive behavior. In Section V, we verify that the obtained results are not uniquely obtained depending on the proposed model and environment but are influenced by the driving behavior. In Section IV, we also interpret the results by comparing them with the knowledge about the visual angle obtained from a previous study [31]. However, to verify the validity of the results regarding the cognitive behavior of individual drivers, we consider it necessary to conduct psychological experiments. For example, in the simulator experiment, we need to confirm that the steering behavior changed when the information in the region with important weights obtained from this study was hidden, or whether the steering behavior did not change when the information in the region with unimportant weights was hidden. Thus, it is necessary to verify in detail the extent to which the machine learning-based approach for driver analysis can be used generally.

VII. CONCLUSION

This study aims to develop a machine learning approach to analyze a driver's cognitive behavior for steering control and to validate that machine learning methods can be an analytical method for understanding driver behavior. We propose a CNN with human physical characteristics and analyze the driving data generated in the simulator study. The results show that the feature maps that are strongly associated with the steering performance are concentrated around the gaze point. The results suggest that drivers perform the steering action using the information within approximately 20 degrees from the gaze point. In particular, the area that more strongly affected the steering is within 10 degrees. This shows that the results obtained from our proposed method can reproduce the same results as previous findings. Next, we confirm that these results are dependent not only on the proposed method and environment, but also on the characteristics of the participants. We analyze the dataset generated by the driver model, which is based on the different driving behaviors of the driver. In the section where we validate this, the results generated by the driver model show a different trend from the results of the human data. Therefore, this validation result shows that the results obtained from the proposed model can indeed reflect the characteristics of drivers. Consequently, these results imply that machine learning methods have the potential to become analytical methods for understanding driver behavior.

The results of this study show the possibility that machine learning methods can be used as a new approach to analyze human cognitive behavior. In previous studies, cognitive behaviors for operations are analyzed by intentionally hiding information in a simulator environment [8], [19], [20], [22]. However, these methods have various limitations, such as the inability to use similar experimental methods in real environments. Whereas, the present approach does not require such a special environment. Additionally, the present study shows the results reflecting cognitive behavior that depends on the driving behavior. Therefore, this study shows the possibility of a new analytical approach to psychological experiments, and is expected to overcome the limitations of conventional psychological experiments. Moreover, although this study focused on the analysis of cognitive behavior of drivers, the machine learning-based method of analysis can be applied to the analysis of not only driving behavior as a whole, but also to the analysis of general-purpose human characteristics that are not confined to driving characteristics.

REFERENCES

[1] T. Bellet, B. Bailly-Asuni, P. Mayenobe, and A. Banet, ''A theoretical and methodological framework for studying and modelling drivers' mental representations,'' *Saf. Sci.*, vol. 47, no. 9, pp. 1205–1221, Nov. 2009.

[2] O. Lappi, ''Future path and tangent point models in the visual control of locomotion in curve driving,'' *J. Vis.*, vol. 14, no. 12, pp. 1–22, 2014.

[3] O. Lappi, E. Lehtonen, J. Pekkanen, and T. Itkonen, ''Beyond the tangent point: Gaze targets in naturalistic driving,'' *J. Vis.*, vol. 13, pp. 1–18, Nov. 2013.

[4] T. Itkonen, J. Pekkanen, and O. Lappi, ''Driver gaze behavior is different in normal curve driving and when looking at the tangent point,'' *PLoS ONE*, vol. 10, no. 8, Aug. 2015, Art. no. e0135505.

[5] G. Li, Y. Wang, F. Zhu, X. Sui, N. Wang, X. Qu, and P. Green, ''Drivers' visual scanning behavior at signalized and unsignalized intersections: A naturalistic driving study in China,'' *J. Saf. Res.*, vol. 71, pp. 219–229, Dec. 2019.

[6] C. J. Nash, D. J. Cole, and R. S. Bigler, ''A review of human sensory dynamics for application to models of driver steering and speed control,'' *Biol. Cybern.*, vol. 110, nos. 2–3, pp. 91–116, Jun. 2016.

[7] D. T. McRuer, R. W. Allen, D. H. Weir, and R. H. Klein, ''New results in driver steering control models,'' *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 19, no. 4, pp. 381–397, Aug. 1977.

[8] D. D. Salvucci and R. Gray, ''A two-point visual control model of steering,'' *Perception*, vol. 33, no. 10, pp. 1233–1248, Oct. 2004.

[9] G. Li, Y. Chen, D. Cao, X. Qu, B. Cheng, and K. Li, ''Extraction of descriptive driving patterns from driving data using unsupervised algorithms,'' *Mech. Syst. Signal Process.*, vol. 156, Jul. 2021, Art. no. 107589.

[10] C. M. Martinez, M. Heucke, F.-Y. Wang, B. Gao, and D. Cao, ''Driving style recognition for intelligent vehicle control and advanced driver assistance: A survey,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 666–676, Mar. 2018.

[11] E. R. Hoffmann and P. N. Joubert, ''The effect of changes in some vehicle handling variables on driver steering performance,'' *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 8, no. 3, pp. 245–263, Jun. 1966.

[12] D. D. Salvucci and A. Liu, ''The time course of a lane change: Driver control and eye-movement behavior,'' *Transp. Res. F, Traffic Psychol. Behav.*, vol. 5, no. 2, pp. 123–132, Jun. 2002.

[13] P. M. van Leeuwen, S. de Groot, R. Happee, and J. C. F. de Winter, ''Differences between racing and non-racing drivers: A simulator study using eye-tracking,'' *PLoS ONE*, vol. 12, no. 11, Nov. 2017, Art. no. e0186871.

[14] B. Wolfe, J. Dobres, R. Rosenholtz, and B. Reimer, ''More than the useful field: Considering peripheral vision in driving,'' *Appl. Ergonom.*, vol. 65, pp. 316–325, Nov. 2017.

[15] B. Reimer, B. Mehler, Y. Wang, and J. F. Coughlin, ''A field study on the impact of variations in short-term memory demands on drivers' visual attention and driving performance across three age groups,'' *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 54, no. 3, pp. 454–468, Jun. 2012.

[16] S. Lemonnier, L. Désiré, R. Brémond, and T. Baccino, ''Drivers' visual attention: A field study at intersections,'' *Transp. Res. F, Traffic Psychol. Behav.*, vol. 69, pp. 206–221, Feb. 2020.

[17] S. Tuhkanen, J. Pekkanen, E. Lehtonen, and O. Lappi, ''Effects of an active visuomotor steering task on covert attention,'' *J. Eye Movement Res.*, vol. 12, no. 3, pp. 1–19, Aug. 2019.

[18] R. Wilkie and J. Wann, ''Controlling steering and judging heading: Retinal flow, visual direction, and extraretinal information,'' *J. Exp. Psychol., Hum. Perception Perform.*, vol. 29, no. 2, pp. 363–378, 2003.

[19] G. K. Kountouriotis, C. D. Mole, N. Merat, and R. M. Wilkie, ''The need for speed: Global optic flow speed influences steering,'' *Roy. Soc. Open Sci.*, vol. 3, no. 5, May 2016, Art. no. 160096.

[20] M. Land and J. Horwood, ''Which parts of the road guide steering?'' *Nature*, vol. 377, no. 6547, pp. 339–340, Sep. 1995.

[21] A. Chatziastros, G. Wallis, and H. Bülthoff, ''The effect of field of view and surface texture on driver steering performance,'' in *Vision Vehicles*, A. G. Gale Ed. Amsterdam, The Netherlands: Elsevier, 1999, pp. 253–260.

[22] Y. Okafuji, C. D. Mole, N. Merat, T. Fukao, Y. Yokokohji, H. Inou, and R. M. Wilkie, ''Steering bends and changing lanes: The impact of optic flow and road edges on two point steering control,'' *J. Vis.*, vol. 18, no. 14, pp. 1–19, 2018.

[23] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, ''Residual attention network for image classification,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3156–3164.

[24] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, ''End to end learning for self-driving cars,'' 2016, *arXiv:1604.07316*. [Online]. Available: http://arxiv.org/abs/1604.07316

[25] M. Bojarski, P. Yeres, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, and U. Müller, ''Explaining how a deep neural network trained with end-to-end learning steers a car,'' 2017, *arXiv:1704.07911*. [Online]. Available: http://arxiv.org/abs/1704.07911

[26] J. Kim and J. Canny, ''Interpretable learning for self-driving cars by visualizing causal attention,'' 2017, *arXiv:1703.10631*. [Online]. Available: http://arxiv.org/abs/1703.10631

[27] Y. Xia, D. Zhang, J. Kim, K. Nakayama, K. Zipser, and D. Whitney, ''Predicting driver attention in critical situations,'' 2017, *arXiv:1711.06406*. [Online]. Available: http://arxiv.org/abs/1711.06406

[28] A. Pal, S. Mondal, and H. I. Christensen, ''Looking at the right stuff: Guided semantic-gaze for autonomous driving,'' 2019, *arXiv:1911.10455*. [Online]. Available: http://arxiv.org/abs/1911.10455

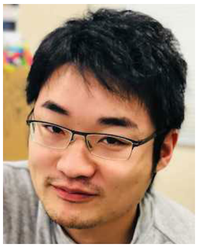
[29] Y. Okafuji, T. Sugiura, and T. Wada, ''Preliminary investigation of visual information influencing driver's steering control based on CNN,'' in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2020, pp. 2263–2268.

[30] L. Saleh, P. Chevrel, F. Mars, J.-F. Lafay, and F. Claveau, ''Human-like cybernetic driver model for lane keepin,'' in *Proc. 18th IFAC World Congr.*, vol. 44, no. 1, pp. 4323–4368, 2011.

[31] H. Strasburger, I. Rentschler, and M. Juettner, "Peripheral vision and pattern recognition: A review," *J. Vis.*, vol. 11, no. 5, pp. 1–82, May 2011.
[32] Y.-Y. Ren, X.-S. Li, X.-L. Zheng, Z. Li, and Q.-C. Zhao, "Analysis of drivers' eye-movement characteristics when driving around curves," *Discrete Dyn. Nature Soc.*, vol. 2015, Mar. 2015, Art. no. 462792.
[33] Y. Okafuji and T. Fukao, "Theoretical interpretation of drivers' gaze strategy influenced by optical flow," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, Art. no. 2389.
[34] K. Tobii. (2017). *Eye Tracker Data Quality Report: Accuracy, Precision and Detected Gaze Under Optimal Conditions-Controlled Environment*. [Online]. Available: https://bit.ly/36nSGqc
[35] Y. Okafuji, T. Wada, T. Sugiura, K. Murakami, and H. Ishida, "Drivers' gaze behaviors are influenced by vehicle position," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 64, no. 1, pp. 1625–1629, Dec. 2020.
[36] G. K. Kountouriotis, R. C. Floyd, P. H. Gardner, N. Merat, and R. M. Wilkie, "The role of gaze and road edge information during high-speed locomotion," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 38, no. 3, pp. 682–702, 2012.
[37] C. D. Mole, G. Kountouriotis, J. Billington, and R. M. Wilkie, "Optic flow speed modulates guidance level control: New insights into two-level steering," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 42, no. 11, pp. 1818–1838, Nov. 2016.
[38] H. Antonson, S. Márdh, M. Wiklund, and G. Blomqvist, "Effect of surrounding landscape on driving behaviour: A driving simulator study," *J. Environ. Psychol.*, vol. 29, no. 4, pp. 493–502, Dec. 2009.

YUKI OKAFUJI (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Kobe University, Hyogo, Japan, in 2014, 2015, and 2018, respectively. In 2018, he was a JSPS Research Fellow (DC2) at Kobe University. From 2018 to 2019, he was a JSPS Research Fellow (PD) at Ritsumeikan University, Shiga, Japan. In 2017, he was a Visiting Researcher with the University of Leeds, Leeds, U.K. Since 2019, he has been an Assistant Professor with Ritsumeikan University. His main research interests include analyzing driver behavior, human modeling, and human–robot interaction.

TOSHIHITO SUGIURA received the B.E. and M.E. degrees from Ritsumeikan University, Shiga, Japan, in 2018 and 2020, respectively. His main research interests include analyzing driver behavior and human modeling.



RYUSEI OSUGI received the B.E. and M.E. degrees from Ritsumeikan University, Shiga, Japan, in 2018 and 2021, respectively. His main research interests include analyzing driver behavior and underwater robot control.

CHENKAI ZHANG received the B.E. and B.A. degrees from the Dalian University of Technology, Dalian, China, in 2019, and the B.E. degree from Ritsumeikan University, Shiga, Japan, in 2019. His main research interests include explainable artificial intelligence and reliability of automatic driving.

TAKAHIRO WADA (Member, IEEE) received the B.S. degree in mechanical engineering, the M.S. degree in information science and systems engineering, and the Ph.D. degree in robotics from Ritsumeikan University, Japan, 1994, 1996, and 1999, respectively. In 1999, he became an Assistant Professor at Ritsumeikan University. In 2000, he joined Kagawa University, Takamatsu, Japan, as an Assistant Professor with the Department of Intelligent Mechanical Systems Engineering, as a Faculty of Engineering. He was promoted to an Associate Professor, in 2003. In 2006 and 2007, he spent half a year at the University of Michigan Transportation Research Institute, Ann Arbor, as a Visiting Researcher. In 2012, he joined the College of Information Science and Engineering, Ritsumeikan University, as a Full Professor. Since 2021, he has been a Full Professor with the Nara Institute of Science and Technology (NAIST). His current research interests include human machine systems, robotics, and human modeling. He is a member of ITSS, SMC, EMBS, RAS, the Robotics Society of Japan (RSJ), the Society of Automotive Engineers of Japan (JSAE), and the Society of Instrument and Control Engineers (SICE). He received the Best Paper Award from JSAE, in 2008 and 2011, the Outstanding Oral Presentation from the Society for Automotive Engineers (SAE), in 2010, and among other awards.