Classifying Cognitive Load and Driving Situation with Machine Learning

Yutaka Yoshida, Hayato Ohwada, Fumio Mizoguchi, and Hirotoshi Iwasaki

Abstract—**This paper classifies a driver's cognitive state in real driving situations to improve the in-vehicle information service that judges a user's cognitive load and driving situation. We measure the driver's eye movement and collect driving sensor data such as braking, acceleration, and steering angles that are used to classify the driver's state. A set of data about the driver's degree of cognitive load, regarded as a training set, is obtained from steering operation and task cognition. Given such information, we use a machine-learning method to classify the driver's cognitive load. We achieved reasonable accuracy in certain driving situations in which the driver moves abnormally for an appropriate service supporting safe driving.**

Index Terms—**Driver's cognitive load, eye movement, machine learning, driving task.**

I. INTRODUCTION

In recent years, a car's value has come to be determined by the new standard of "information performance." Vigor is a current car-navigation system that will ship about four million sets in 2006 and surpasses an 80% installation rate in new cars. Vigor is becoming indispensable to many users for a comfortable and convenient drive. Still, the possibility of a traffic accident due to an in-vehicle information service also exists. A driver's car-navigation system operation and its audio operation could cause an automobile accident. Even if it is not the direct cause, we surmise that the attention given to the driving task is decreased when a driver pays attention to in- vehicle information or a news bulletin. Alternatively, the user may miss the useful information that an in-vehicle information service offers while concentrating on the driving task.

As mentioned above, we need a function within the in-vehicle information service that judges the user's cognitive load and driving situation. The driving situation determines to what degree the user should concentrate on the driving task. The purpose of this research is to judge a user's cognitive load and driving situation using sensor information from an in-vehicle system.

This paper focuses on how a user's cognitive load and driving situation can be determined and classified using features involving vital reactions (eye movement) and user operations (driving the car). For example, a driver is cautious about a person in a pedestrian crossing when turning left. In this case, the cognitive load is increased by a change in the driving environment. Moreover, in order to recognize the

surrounding situation, the range of eye movement may increase because the driver's cognitive load has increased. This means that a high cognitive load and a change of eye movement are related. Eye movement is used in the field of physiological psychology for clarifying control [1]. It is directly related to perception and can be considered an indication of mental load. Driving a car requires prediction of the surrounding environment and is influenced by the situation. Therefore, a user's cognitive load can be classified using these features. To do this, we measure the driver's eye movement and gather driving data such as accelerator use, braking, and steering.

This paper takes a machine-learning approach to the above cognitive state identification problem in a realistic car-driving task. We set up cognitive loads according to the steering-entropy method [2] and a definition of the task cognition situation. This paper is organized as follows. Section II presents related works. Section III defines cognitive load. Section IV describes our classification model for cognitive load. Section V presents a performance evaluation. The final section provides conclusions.

II. RELATED WORKS

Much research exists analyzing mental models, using knowledge of cognitive science or brain science, mathematical equations [3], stochastic models [4], and engineering-control models [5].

As a related work, [6] focuses on the task of lane maintenance and calculates the motion of vehicles using a machine-learning model for vehicle control. [7] constructed a driver's cognition and action model by using ACT-R, which is a cognitive architecture. [8] presented a model of the driver's cognitive load using resource-allocation theory and a qualitative model. These models assume the model beforehand and are concerned with verifying the closeness of fit to the data. [9] used inductive-logic programming (ILP) to learn a model of driving from data, and examined differences between the country and a city. In this paper, we used the same driving data but focused on a measurement of cognitive state that differs from that in the paper [9].

III. COGNITIVE LOAD

First, we explain the concept of cognition proposed so far. Perception and cognition are general concepts for driving. Perception is a process in which information is directly taken in from a sense organ. Cognition is a process in which a person understands the information about an external object or phenomenon in response to the influence of acquired knowledge, memory, and experience. These concepts appear

in car-driving textbooks in Japan and are generally used to model car driving [10]. For example, when a car approaches an intersection with poor visibility, the driver treats the intersection as an important thing (perception), and the driver also predicts important phenomenon in which people come from a street intersection (cognition). Moreover, in this paper, phenomena related to accidents are called car-driving risks. The following subsections describe two measurements of cognitive load.

Fig .1. Example of a situation with two or more tasks.

A. Task Cognition

In this research, a driving task is adopted as one of the measurements describing cognitive load. Drivers often cope with two or more tasks at a time when driving. Fig. 1 portrays such a situation. We numbered the objects that the driver should pay attention to in this situation. First, the driver needs to manage the accelerator and brake so that he does not collide with the car in front. Second, a man on a bicycle may go to the front pedestrian crossing or to the left rear. Third, the driver needs to pay attention to the left side because a pedestrian or a motorbike may come from that direction. Moreover, the driver needs to perform lateral control appropriately in order to move forward along a curve. Fourth, this route has a place where cars merge from the right. The driver fully needs to be cautious of cars approaching from the right and needs to move forward. In such a situation, the driver's reaction to this new risk is delayed. This research considers such a situation to have high cognitive load.

We determine the range of time of the high cognitive load as seen in Fig. 2. Fig. 2 depicts the time relation between perception, task cognition, and driver action. According to the cognitive definition, when a driver perceives a stopped vehicle, the driver predicts a collision with it. This cognition is continued until the task of avoiding a stopped vehicle ends or the stopped vehicle disappears from the field of view. We define this as the task-cognition time.

B. Steering Entropy Method

The steering-entropy method is used as the second measurement of cognitive load [2]. The steering-entropy method quantifies the roughness of steering when a driver operates under various loads (i.e. listening to traffic information from the radio, talking, doing mental arithmetic, etc.).

The theory underlying the steering-entropy method was developed in order to investigate the degree to which the operating car equipment burdens the driver. We consider that this theory is similar to the cognitive-load approach taken in our research. Sometimes a person focuses too much on a certain object or task when driving a car and neglects other factors. This theory can thus measure a driver's cognitive load.
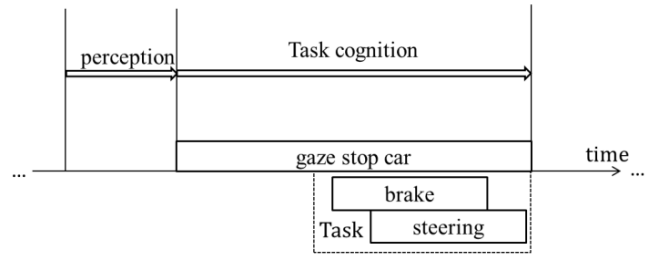

Fig. 2. The time range of task cognition.

The human's shortest control interval is 150ms and measures the error of the steering angle with this time interval. In [2], the task is performed for 6 to 8 seconds, and the increase of steering entropy is measured. In this research, steering entropy is measured at intervals of 10 seconds.

Let $\theta(t)$ be the steering angle at time t. Using the steering angles of the previous three time steps ($\theta(t-3)$, $\theta(t-2)$ and $\theta(t-1)$), we perform a second-order Taylor expansion on time $t–1$ to obtain the predicted steering angle $\hat{\theta}(t)$ at time t (i.e., the steering angle likely to be obtained if steering is executed very smoothly). This is calculated using Eq. (1).

$$\hat{\theta}(t) = \theta(t-1) + (\theta(t-1) - \theta(t-2)) + \frac{1}{2}\{(\theta(t-1) - \theta(t-2)) - (\theta(t-2) - \theta(t-3))\} \tag{1}$$

Prediction error $e(t)$ is defined as the difference between $\hat{\theta}(t)$ and $\theta(t)$.

$$e(t) = \hat{\theta}(t) - \theta(t) \tag{2}$$

The 90th percentile value α of the frequency distribution of the recorded prediction errors is then computed. The prediction error distribution becomes narrower (smaller) as the driver's steering becomes smoother. Parameter α indicates the fundamental steering response of an individual. The frequency distribution is then divided into nine bins based on this α (the borders are at -5α, -2.5α, $-\alpha$, -0.5α, 0.5α, α, 2.5α, and 5α). The proportion of prediction errors $p_1, p_2, \cdots, p_9$ falling into each bin is computed, and steering entropy H_p is calculated using Eq. (3).

$$H_p = -\sum_{i=1}^{9} p_i \log_9 p_i \tag{3}$$

IV. CLASSIFICATION MODEL OF COGNITIVE LOAD

This section explains the features used to classify the

cognitive load discussed in Section III. We use the car's state, the car, the driver's actions, eye movement, and gaze targets as features.

TABLE I: CAR'S STATE AND DRIVER'S ACTION

Longitudinal control	accelerator+
	accelerator-
	brake+
	brake-
	accelerator to brake
	brake to accelerator
Lateral control	right
	stationary
	left
Gear selection	1 up
	stationary
	1 down
	2 down
Gaze target	rearview mirror
	right mirror
	left mirror
	car navigation
	speedometer
	right viewing
	left viewing
	duration of gaze
Car's state	speed
	rate of acceleration
	accelerator rate
	steering angle
	gear number
	change of azimuth direction

TABLE II: EYE MOVEMENT AND THE TARGETS FOR GAZING

Eye movement feature	variance of a look position (*X, Y* coordinate)
	variance of degree of the angle of a look moving direction (radian)
The target for gazing	traffic light
	road sign
	pedestrian
	vehicle in front
	oncoming vehicle
	standing vehicle
	direction lane
	other lane
	oncoming lane
	corner
	curved mirror
	sidewalk
	pedestrian crossing
Attribute of the target for gazing	average of *X* coordinate
	average of *Y* coordinate
	continuation time of a gaze

A. Car's State and Driver's Action

A car is driven through the interaction of driving actions and the car's state car. When the driver's cognitive load increases, the car's state and the driver's action differ from a situation in which the driver's cognitive load is normal. For example, when it slows, the driver steps on the accelerator and increases the car's speed. The action depends on the driver's situation. If it is a straight road with high visibility, a driver may accelerate even if the speed is already high. Conversely, if it is a bad road with many curves, the driver may decelerate even if the speed is low. Furthermore, if the driver's cognitive load is high, the accelerator may be disengaged even if the speed is low on a good road. As mentioned above, a driver's action depends on the car's state or the driver's cognitive condition. We can thus classify a driver's cognitive condition from the combined car's state and the driving action. The driving actions and car states used in this research are listed in a Table I.

A "longitudinal control" action is an action that manages the forward motion of a car, which here means operation of the accelerator or brake. When the accelerator rate exceeds zero and the change is positive, it was considered as "accelerator+." When the accelerator rate exceeds zero and the change is negative, or when the foot was not moved to the brake but the accelerator rate was zero, it was considered as "accelerator-." It was considered "brake+" when a brake signal was detected, and it was considered "brake-" when the foot was not moved to the accelerator and no brake signal was detected. a brake signal was not detected. A movement to the brake from the accelerator is considered "from accelerator to brake" when a brake signal was detected after the acceleration became zero. A movement to the accelerator from the brake is set considered "from brake to accelerator" when the acceleration was positive and a brake signal was not detected.

A lateral-control action is an action that manages the direction of the car, which here means steering wheel operation. We divided this into three classes, left change (left), right change (right), and stationary, based on the change of the steering-angle signal.

Although shifting gears is essentially a longitudinal control, it was considered an exceptional action since the gearshift is located away from the accelerator and the brake. The gear operation was divided into four classes, one increase (1 up), one reduction (1 down), two reductions (2 down), and stationary, based on changes in the gear signal. Attentive actions include gazing at side mirrors, the rearview mirror, the speedometer, car navigation, and right-and-left viewing. Unlike other targets for gazing, these are attached to the car, and can be seen at any time. Since these acts are performed actively so a driver can understand the outside environment, these are considered to be driver actions.

B. Eye Movement and Gazing Targets

In this research, the driver's eye movement and the gazing targets are used for classification. We measured a driver's eye movement by using an eye movement tracking device. The features involving eye movement and gazing targets are summarized in Table II. The features of eye movement use the variance of a looking position (*X*, *Y* coordinates), and the variance of the angle in the looking direction (radians). A motion of the eyeball is faster quicker than the control of one's hand or leg. Thus, the steering entropy was computed at 10-second intervals, but the variance of eye movement was computed at five-second intervals. We surmise that the variance of a looking position and the variance of the angle in the looking direction differ depending on the car's state or the driver's cognitive load. For example, when objects to be avoided (e.g., a pedestrian or stopped cars) exist on both the right and left sides, the variance of eye movement may become large, and the cognitive load may be high. The

gazing target represents the driver's recognition of the outside environment.
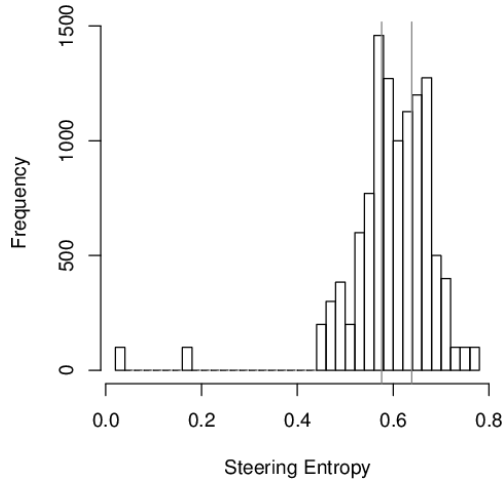

Fig. 3. Histogram of steering entropy.

It is useful to consider the gaze position (*X, Y* coordinates) and the gaze duration. There is a time lag between gazing upon an object taking action. For example, if the object to be avoided is recognized from a long distance, the driving risk is recognized at an early point and the object can thus be avoided in a smooth operation. We assume that the driver's cognitive load would be small at this time.

C. Classification Model

This section explains the model used to classify a driver's cognitive condition. If a driver's cognitive condition can be predicted in real time, the in-vehicle system can provide service according to the cognitive state. In this paper, steering entropy and task cognition are treated as measures of cognitive state and are predicted. We used two machine-learning methods for the classification. We used the "support vector machine" (SVM), which is a learning models with excellent pattern recognition, and the "random forest" (RF), which is a decision-tree model of ensemble learning. We used the features in Tables I and II as feature vectors.

Since we are interested in the magnitude of the steering entropy rather than the exact value, we performed "quantile discretization" [11]. This method discretizes the data so that there are the same numbers of samples in each category after discretization. A scene that does not present a high cognitive load belongs to the high cognitive load class if the steering entropy is divided into two classes. We therefore divide the steering entropies into three ranges. The division range is seen in Fig. 3. The boundary between the small and middle class is 0.576, and the boundary between the middle and large class is 0.639. We convert this into a discrete variable that is set to 1 (positive example) for a large range (exceeding 0.639), and set to 0 (negative example) for a minor range (0.639 or less). The task cognition was set to 1 (positive example) when it existed, and set to 0 (negative example) when it did not exist.

V. EXPERIMENTAL EVALUATION

A. Dataset

We used an eye movement tracking device. The device can measure horizontal and vertical viewing angles in degrees.

We obtained 60 data points per second.

TABLE III: THE INFORMATION OF EACH INTERVAL

Interval number	Time (0.1s)	Number of steering entropy positive examples	Number of task cognition positive examples
1	1002	200	579
2	576	176	355
3	150	100	89
4	2427	1000	531
5	1432	600	125
6	326	126	114
7	771	400	222
8	1362	600	240
9	852	100	73
10	1735	300	635

TABLE IV: ABOUT OF TP, TN, FP, FN

		observed	
		1	0
predicted	1	TP	FP
	0	FN	TN

A Controller Area Network (CAN) is an in-vehicle LAN used to gather driving data. We can obtain the accelerator depression rate (0% to 100%), braking signal (0 to 1), steering signal (-450 to 450 degrees), a signal representing the gear (0 to 4), vehicle speed (km/h), azimuth direction (0-360 degrees), and so on. To measure these, we modified a Toyota Crown and obtained 10 data points per second. The azimuth direction was measured at one data point per second so we smoothly interpolated the data using a spline function [12].

Furthermore, we manually labeled the gazing targets based on eye-movement data. The labels were assigned 10 data points per second. The eye-movement variance data was measured every five seconds. We manually evaluated and labeled task cognition every 0.1 second based on the animation data, according to the definition in Section 3. The steering entropy was measured every 10 seconds. The total recording time for the data was 25 minutes and 52 seconds. A car speed of zero is included in the data. We assume that the features of driving actions and eye movement differ between a stopped car and a moving car. Therefore, we exclude the data in which the car is stopped. The data are divided into 10 situations each in which the car's speed exceeded 0. We evaluate the classification model by cross-validation for each interval. First, we choose one interval to be used as test data. Second, we generate a classification model by using all of the other intervals as training data. Finally, we investigate the classification accuracy in the one chosen interval. This process is repeated 10 times. The time length, the number of positive examples of steering entropy, and the number of positive examples of task cognition of each interval are listed in Table III.

B. Evaluation of the Classification Model

We used SVM software "kernlab" [13] which is a package of *R*. We used "C-SVM" for the classification and the kernel used the "RBF kernel." We set the cost parameter to 0.1, 1, 10, and 100. The cost parameter is order of 0.1, 1, 10, and 100. Cost parameter is the size of the size of the penal regulations

of the error in training. The larger it is, the more it performs over-fitting. We used random forest software "randomForest" [14], which is also a package of *R*.
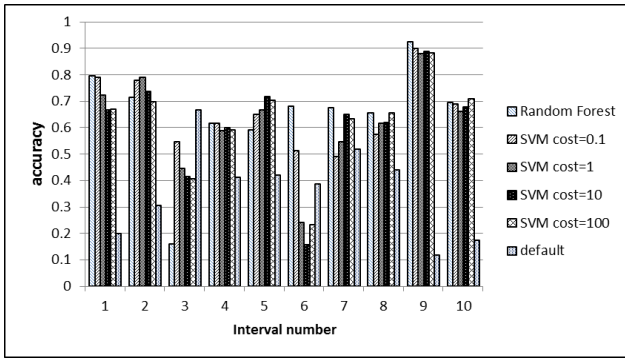

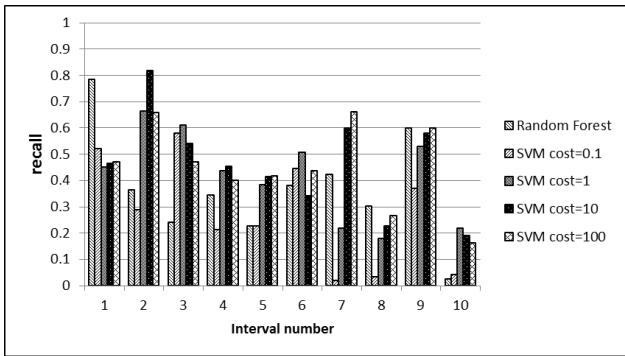Fig. 4. Classification accuracy of the steering entropy.

Fig. 5. Classification recall of the steering entropy.

We use accuracy and recall as evaluation measurements for classifying a binary variable. The accuracy was calculated by using Eq. (4), and the recall was calculated by using Eq. (5). TP, TN, FP, and FN in the formula respectively represent True Positive, True Negative, False Positive, and False Negative, and these are related as seen in Table IV.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$recall = \frac{TP}{TP+FN} \quad (5)$$

The default accuracy is the predictive accuracy when all of the data is classified into a positive example.

The classification accuracy of the steering entropy in each of the machine-learning methods is illustrated in Fig. 4. The classification recall of the steering entropy in each of the machine-learning methods is illustrated in Fig. 5. The classification accuracy of task cognition in each of the machine-learning methods is illustrated in Fig. 6. The classification recall of the task cognition in each of the machine-learning methods is illustrated in Fig. 7.

VI. Discussion

Regarding the prediction of steering entropy, the predictive accuracy produced a greatly different result for every interval. Each of the intervals except 3, 6, and 7 yielded a result better than the default accuracy. In some intervals, the cost of SVM is lower to perform with high accuracy; in other intervals, the cost of SVM is higher to perform with high accuracy. This demonstrates that there no SVM value is optimal for all intervals. However, the recall of a cost that exceeds one is generally higher than a cost that is 0.1. Since the cognitive load should be detectable, a parameter cost of one or more is good.
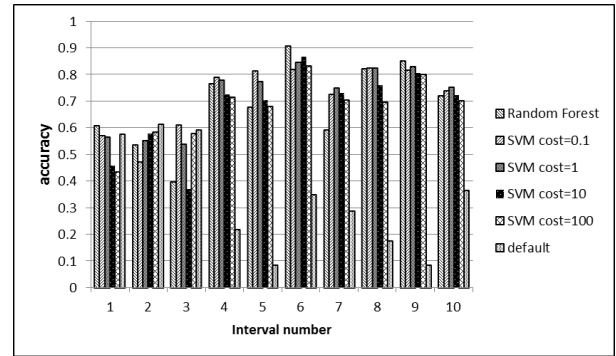

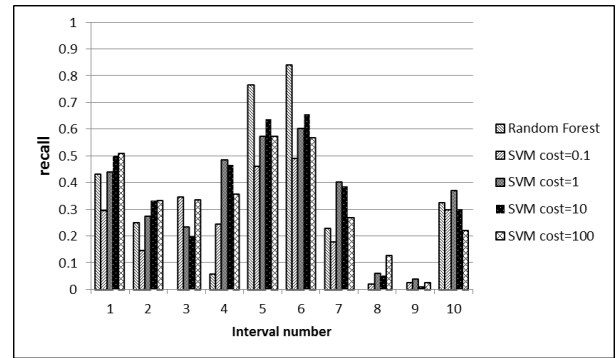Fig. 6. Classification accuracy of task cognition.

Fig. 7. Classification recall of task cognition.

Regarding the task cognition prediction, intervals other than intervals 1 to 3 have a predictive accuracy better than the default accuracy. However, the recall of intervals 8 and 9 is almost zero and cannot detect task cognition. In contrast, intervals 5 and 6 had comparatively high recall, and their predictive accuracy was also good. Intervals 5 and 6 having sufficient accuracy and recall because there are many curve tasks, brakes are applied, and significant turning of the steering wheel is necessary in many cases in order to avoid an object (a pedestrian, stopped vehicle, etc.). In contrast, the recall of intervals 8 and 9 is near zero because the driver's action is not reflected strongly, even though the driver recognizes the task. The driver generally did not slow down when entering at a curve. The driver has not taken evasive action (stepping on the brake or turning the steering wheel), even when passing an oncoming car on a narrow road. The accuracy of task cognition detection is good when the driver performs an action accompanied by a significant change. The feature representing gazing did not contribute to the classification other than with regard to seeing a curve or a pedestrian. There is thus no difference in the target attention time and the target for gazing between a normal driving situation and a task recognition situation.

VII. Conclusion

In this paper, we constructed a classification model of the cognitive load on a human being when driving a car. If a

cognitive state can be classified based on a feature of the driver's action and the car driving environment obtained from an in-vehicle system, the system can recognize a driver's situation and can automatically operate an apparatus. We predict that the classification model of a cognitive state will contribute to realizing a new in-vehicle system that provides automatic support according to a driver's cognitive condition.

We used the steering task cognition and entropy method as measurements of the driver's cognitive load. First, we labeled task cognition manually when the driver recognizes the driving task according to the driver's perception and the cognitive definition. Second, the steering-entropy method evaluated the roughness of steering-angle operation for tasks other than driving actions. Since we assume that this is similar to the cognitive load considered by this research, a driver's cognitive load is expressed based on this theory.

We used the information about the driver's actions, the state of the car, eye movement, and the gazing target as features used by the classification model. We used "SVM," a learning model with excellent pattern-recognition performance, and "random forest," a decision-tree model of ensemble learning.

The classification demonstrated that steering entropy and task cognition could be predicted except for certain intervals. The accuracy of task cognition detection is good when the driver performs an action accompanied by a significant vehicle response.

In future work, we will investigate in detail the influence that the car driving environment has on a driver's cognitive state, and the influence that cognition has on a driver's actions or vital reactions.

References

[1] R. H. S. Carpenter, *Movements of the Eyes*, Pion Ltd.1988.
[2] O. Nakayama, T. Futami, T. Nakamura, and E. R. Boer, "Development of a steering entropy method for evaluating driver workload," *International Congress and Exposition*, 1999
[3] J. Enderle, S. M. Blanchard, and J. Bronzino, *Introduction to Biomedical Engineering*, Academic Press, 2005.
[4] J. K. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. L. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex," *Nature Neuroscience*, vol. 2, no. 7, pp. 664-670, 1999.
[5] J. M. Findlay and R. Walker, "A model of saccade generation based on parallel processing and competitive inhibition," *Behavioral and Brain Science*, vol. 22, pp. 661-721,1999.
[6] D. Pomerleau and T. Jochem, "Rapidly adapting machine vision for automated vehicle steering," *IEEE Expert*, pp. 19-27, 1996.
[7] D. D. Salvucci, "Modeling driver behavior in a cognitive architecture," *Human Factors*, vol. 48, no. 2, pp. 362-380, 2006.
[8] S. Sega, H. Iwasaki, H. Hiraishi, and F. Mizoguchi, "Qualitative reasoning approach to the driver's cognitive mental load," *International Journal on Software Scienie and Computational Intelligence*, pp. 18-32, 2004 (in press).
[9] F. Mizoguchi, H. Ohwada, H. Nishiyama, and H. Iwasaki, "Driver's cognitive load identification using inductive logic programming," *Inductive Logic Programming Lecture Notes in Computer Science*, vol. 7842, pp 166-177, 2013.
[10] E. Hollnagel and D. D.Woods, "Joint Cognitive Systems:Foundations of Cognitive Systems Engineering," pp. 149-154, CRC Press, 2005.
[11] A. Hartemink, "Principled computational methods for the validation and discovery of genetic regulatory networks," Ph.D. thesis, School of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2001.
[12] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, "computer methods for mathematical computations," 1977.
[13] A. Karatzoglou, A. Smola, and K. Hornik, "Kernlab – An S4 Package for Kernel Methods in R," *Journal of Statistical Software,* pp. 1-20, 2004
[14] L. Breiman, "Random Forest," *Machine Learning*, vol. 45, no. 1, pp.5-32, 2001.

Yutaka Yoshida graduated from the Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science, Noda City, Japan, in 2012.

He is a student at Tokyo University of Science Graduate School, Division of Science and Engineering Industrial Administration Master's course since 2012, Noda City, Japan. His research interests are in the field of machine learning and application.

Hayato Ohwada graduated from the Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science, Noda City, Japan, 1983. Then he graduated from Tokyo University of Science Graduate School, Division of Science and Engineering Industrial Administration Doctoral course completed program with degree in 1988.

He was a research associate from 1988 to 1998, a lecturer from 1999 to 2000, and an associate professor (Tokyo University of Science) from 2001 to 2004. Then he is a professor at Tokyo University of Science Faculty of Science and Engineering from 2005. His research interests are in the fields of Inductive Logic Programming and Bioinformatics.

Fumio Mizoguchi was an research associate in the Tokyo University of Science from 1968 to 1974, a lecturer from 1974 to 1982, an associate professor from 1982 to 1983, and a professor from 1987 to 1989, a research committee from 1989 in the Tokyo University of Science .

Currently, he is visiting professor of research division of the next generation data mining at Tokyo University of Science.

Hirotoshi Iwasaki graduated from Department of Electrical Engineering, Faculty of Technology, Nagoya University, 1988. Then he graduated from Nagoya University Graduate School, Division of Electrical and Electronic Engineering Master's Program with degree in 1990. Currently, he is a researcher of Denso IT Laboratory, Inc., Japan .