IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 23, 2018, accepted November 2, 2018, date of publication November 6, 2018,
date of current version December 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2879887

Modeling Driver Risk Perception on City Roads Using Deep Learning

PENG PING[1,2], YUAN SHENG[2], WENHU QIN[1], CHIYOMI MIYAJIMA[3], (Member, IEEE), and KAZUYA TAKEDA[2], (Senior Member, IEEE)

[1]School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China
[2]School of Informatics, Nagoya University, Nagoya 464-0814, Japan
[3]School of Informatics, Daido University, Nagoya 457-8530, Japan

Corresponding author: Wenhu Qin (qinwenhu@seu.edu.cn)

This work was supported in part by a Grant-in-Aid for Scientific Research (C) under Grant 15K002231 from the Japan Society for the Promotion of Science (JSPS), in part by the National Natural Science Foundation of China (Grant 61300101), and in part by the Key Research Plan of Jiangsu Province (BE2017035).

ABSTRACT Research on how risk is perceived by drivers is vital to driving behavior research and driving safety. As risk can be divided into subjective and objective risk, in this paper, we focus on modeling subjective risk perception by drivers using a deep learning method. Different drivers often perceive different levels of subjective risk under the same driving conditions. In addition, different driving conditions or driving events will have different effects on drivers. Based on these two risk perception features, in this paper, we first design an experiment on a city road with two lanes to assess the level of subjective risk perceived by drivers belonging to different groups. We then use a deep learning network-based method to abstract features of the driving environment. These environmental features are integrated with driver risk perception data and this information is used as training and testing data for the learning network. Finally, a long–short-term memory-based method is adopted to model the subjective risk perception of individual drivers based on traffic conditions and vehicle operation data from the driver's vehicle. Our results show that the proposed method can effectively model the subjective risk perception behavior of drivers, allowing for end-to-end risk perception prediction in future driving assistance systems.

INDEX TERMS Driving behavior modeling, deep learning, traffic safety, risk perception.

I. INTRODUCTION

According to the 2015 traffic safety report from the World Health Organization, 1.2 million lives are lost annually as a result of traffic accidents [1]. Traffic accidents are also the leading cause of death for people between 15 and 29 years of age. Although the annual number of traffic deaths has plateaued at 1.25 million in recent years, justifying the 50% reduction goal by the 2020s is still far away. Unsafe driving behavior is the main cause of traffic accidents, and is responsible for 90% of all fatal crashes [2]. Specifically, distorted and faulty perception of potential risk is one of the main causes of road accidents [3]. Research by M.S. Horswill and F.P. McKenna suggests that among driving skills, only hazard perception is correlated with drivers' accident histories [4].

Risk perception while driving is the subjective assessment of the probability of a specified type of accident and how concerned drivers are with the possible consequences [5]. As it is based on judgement and situational awareness, risk perception is a type of behavior which varies among individuals [6]. Drivers operate their vehicles based on their perceptions of current traffic conditions or events. Inappropriate risk perception leads drivers to fail to recognize or to ignore dangerous situations, or to adopt risky driving behavior. Therefore, it is vital that we investigate how driver risk perception works, and use our understanding to develop the necessary interventions or improvements for driving assistance systems in order to achieve the goal of improving driving safety.

Perception of traffic risk is highly individualized and is influenced by personal experience with the costs of accidents and the rewards for risk-taking behavior [7]. Risk perception is also correlated with objective driving conditions or events. Therefore, this study aims to build a model which uses environmental factors to simulate a particular driver's risk perception behavior on a city road. As shown in Fig. 1, three steps are proposed to reach this goal. First, a risk perception evaluation experiment is designed and conducted to collect risk perception data from drivers and to verify that risk perception is personalized. Second, objective features of the driving

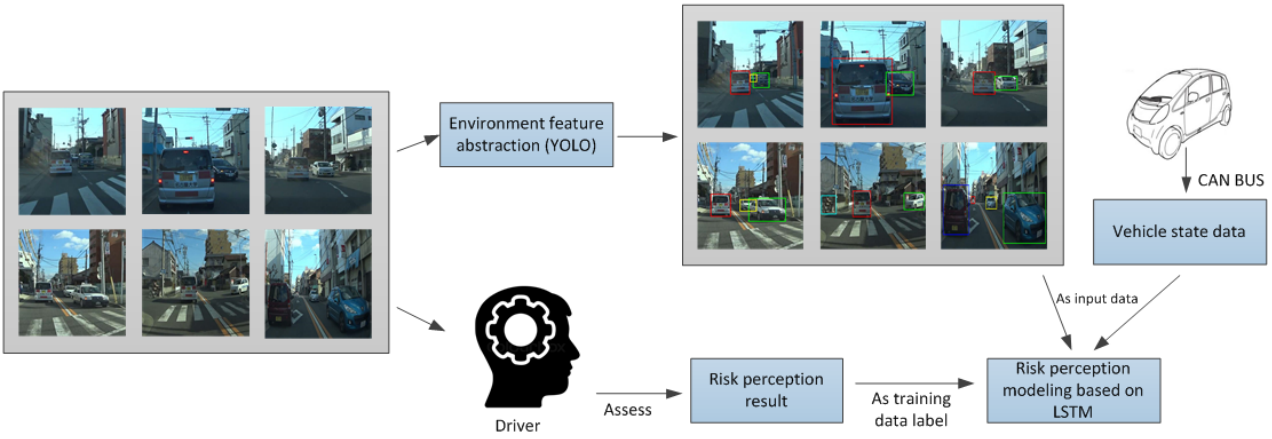


FIGURE 1. Research structure for risk perception modeling.

environment are abstracted and condensed into a usable form which can be compared with driver risk perception. Finally, a risk perception model based on deep learning network is constructed to simulate the risk perception of a particular driver in relation to the driving environment.

Many studies have analyzed driver risk perception, and we can divide the approaches used into three major categories. The first type of these studies is based on questionnaire surveys or self-reported information from drivers, which is effective for illustrating variations in human risk perception. Siren and Kjær conducted textual interviews and group discussions to examine how older drivers conceptualize risk perception [6]. Their results show that older drivers tended to consider risk as something external (dangerous traffic condition or risk driving behavior of other drivers) which can be dealt with by internal means. Rafaely *et al.* conducted a questionnaire survey to examine the difference between older and younger adults in regards to driving safety [8]. Participants in both age groups were able to accurately assess their own risk level when driving, but older participants tended to overestimate the risk incurred by younger adults and younger participants tended to underestimate the risks taken by older adults. Questionnaire surveys are also effective for measuring the effects of some social factors which affect driver risk perception, such as media influence [9], socio-economic level and driving experience [10], cultural influences [11], and so on. Since questionnaire studies reflect the traits of groups or macro knowledge on driving risk perception, they are not as useful when evaluating individuals. In other words, questionnaire results can only reveal the major risk perception features of some groups. Additionally, survey results from questionnaires rely heavily on the samples collected and are easily affected by fluctuation in the survey objects.

The second approach employs driving simulators to collect realistic driver behavior data from drivers encountering high risk situations. By using simulators, the researchers can safely collect risk perception data under realistic driving conditions. Crundall *et al.* analyzed the different responses to risk of a group of experienced drivers [12]. After exposing the drivers to different hazardous traffic scenarios on the driving simulator, the authors proposed that the ability of a driver to perceive risk is based on whether or not the driver can quickly spot possible hazards. Dixit *et al.* used a left turn driving scenario on a driving simulator to analyze the psychological features of subjective risk perception by drivers [13]. They concluded that drivers with better risk perception ability were more optimistic about executing the turn without crash. Experienced drivers also tended to finish the task more safely and efficiently. The use of driving simulators is also valuable when analyzing the effect of specific driver traits on risk perception, such as the responses of elderly drivers when encountering hazardous situations [14], cognitive abilities and the effects of distractions [15], the ability to scan for hazardous objects [16], and so on. Driving simulator-based research also has its shortcomings, however. Although simulators can accurately recreate various driving conditions and scenarios, they cannot reproduce the full complexity of real-world driving conditions, and the simple background and game-like interface also lowers accuracy and realism.

The third method used to study risk perception is to show study participants video clips of real-world driving situations to assess their risk perception abilities. The video clips can be carefully selected to focus on unique, personalized features. Borowsky *et al.* used six selected video clips to compare the hazard perception abilities of young drivers (inexperienced and experienced) with those of elderly drivers [17]. Their study found that older and experienced drivers are more sensitive to risk, confirming the results of previous studies [18], [19]. Crundall used thirty clips of hazardous driving situations and ten of non-hazardous situations to compare the responses of novice and experienced drivers [20]. Gugliotta *et al.* used 300 videos made by two experienced drivers to explore the impact of situational awareness and decision-making skills on hazard perception [21]. Risk perception assessment using real-world video clips allows subjects to experience real driving environments, and because our modeling method is based on environmental factors

which can be visually perceived by drivers, in this study we also use real-world video to assess the risk perception behavior of drivers.

Previous studies have, to a large extent, answered the questions of how the psychological characteristics of various demographic groups affect risk perception, and how different levels of risk perception ability affect traffic safety. In this paper, we want to go a step further and build a quantitative personal risk perception model which can simulate human risk perception, instead focusing on the phenomenon of risk perception itself. There have been a few previous attempts at quantitative modeling of driver risk perception very similar to this study. Liu *et al.* used an adaptive neuro-fuzzy inference system to model driver risk perception, and found that risk prediction results can serve as an important parameter for predicting the intersection crossing behavior of other drivers [22]. The risk perception model was based on a simple intersection scenario which only considered the limited environmental factors of two vehicles crossing an intersection. As a result, their risk perception model cannot be used to model more complex driving environments. Zhao *et al.* utilized fuzzy C-means clustering to examine the driving states experienced by drivers in extreme detail, where each cluster of data represents distinct risk perception behaviors [23]. Risk labelling in this study is done using thresholds of detailed analyses of driving states instead of the cognition of drivers. Although Zhao *et al.* used the micro driving states of drivers to analyze risk perception, they were actually analyzing objective risk instead of subjective risk perceived by drivers.

Risk perception can be treated as a human behavior similar to writing or talking, both of which are the output of human cognition results in response to the surrounding environment. In addition, risk perception is a time sequence behavior which is based on information about previous and current states, most of which is obtained through the channel of vision. Naturally, we want to use some time sequence information processing methods for modeling driver risk perception. Long Short-Term Memory network (LSTM) [24], which is an improved form of Recurrent Neural Network [25], is a powerful pattern recognition tool which is widely used in the modeling of human-like behavior, such as image captioning [26], machine translation [27] and driving behavior modeling [28]. Inspired by these previous research, in this paper we use selected video clips to assess the risk perception ability of drivers, and then use an LSTM-based method to model their risk perception behavior.

The organization of this paper is as follows. Section II presents the methods used to experimentally obtain driver risk perception assessment data and describes our method for abstracting driving environment features. Section III describes the LSTM-based risk perception modeling process. The group and personal modeling results are given in the section IV. Finally, our discussion and conclusion of this study are summarized in Sections V and VI, respectively.

II. METHOD OF DATA COLLECTION

A. RISK PERCEPTION DATA COLLECTION

1) PARTICIPANTS

The participants in our study were 22 Japanese drivers from 24 to 74 years of age, all of whom hold Japanese driving licenses. They were grouped into four categories and a detailed description of each group is given in Table 1.

TABLE 1. Descriptions of driver groups.

Category	# of participants	Age	Description
Elderly drivers	7	from 65-74	Drive more than five times a week, and hold driving licenses for more than 30 years.
Novice drivers	6	from 24-45	Obtained the driving license within two year, or drive less than twice half a year.
Experienced drivers	6	from 43-51	Drive more than five times a week, and hold driving licenses for more than 20 years.
Driving instructors	3	from 32-45	Instructors from a driver training school

2) VEHICLE AND ROAD USED FOR DRIVING VIDEO COLLECTION

The driving data and test videos were collected during real-world driving. A Mitsubishi i-MiEV electric vehicle equipped with a Sony HDR-AS100V camera was used for data collection, as shown in Fig. 2. The camera was placed inside the vehicle and recorded the driving environment from the driver's point of view. The driving route was 5.2 km of two-lane city road (left side driving, one lane traveling in each direction) in Nagoya, Japan. Driving the entire route took

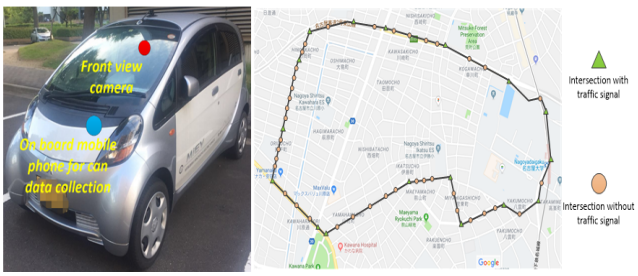


FIGURE 2. Experimental vehicle and data collection route.

about 20 minutes. An on-board mobile phone and video camera recorded vehicle status information and traffic conditions during each driving session. The collection rate was 10Hz, i.e., during each second we obtained 10 samples of CAN data and 10 samples of driving environment feature images (ten video frames). Details about the collected data and its categorization will be introduced in Section II.B.

3) COLLECTION OF DATA FOR RISK PERCEPTION ASSESSMENT

We used the driving videos collected from the real driving of the experiment participants to assess the participants' subjective risk perception behavior. Each of our 22 participants drove along the designated route twice. Videos were collected during the morning, noon and afternoon, and a total of 44 videos (22 drivers$\times$2 trips) were collected for our study. The videos were stored as MPEG-4 files with a resolution of 1920$\times$1080 per frame and then shown to 14 of the participants on a 50-inch TV screen. These 14 participants were asked to assign risk level scores for each frame while viewing the videos, using a software application. The assessment software interface is shown on the left of Fig. 3, and the right side of the figure shows the viewing environment.

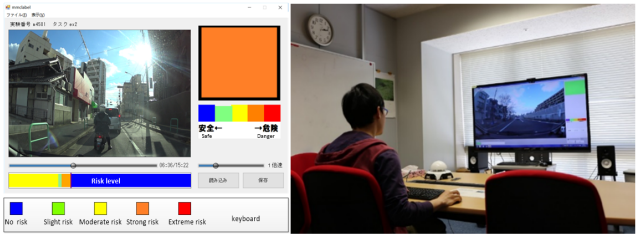


FIGURE 3. Software platform and video viewing environment used in our experiment.

The 14 participants who were asked to evaluate the driving videos were briefed on the research goal of the experiment and then given detailed instructions for using the evaluation software. The participants evaluated the risk levels observed in their own driving videos as well as in those collected by the other 21 participants. As the participants sat in front of the TV screen and watched the videos, they pressed designated keys on a keyboard corresponding to the risk level they perceived at that point in the video. There were five risk level buttons available on the keyboard (no feeling of risk, slight feeling of risk, moderate feeling of risk, strong feeling of risk and extreme feeling of risk), and the software application recorded this risk perception data for each video frame. In our risk assessment process we did not formulate a detailed definition of 'risk' or provide the participants with any guidelines, because our goal was to model subjective risk as perceived by the drivers, rather than to evaluate their risk perception skills. For example, if a participant had poor risk perception skills, our proposed learning network would need to simulate his or her poor risk perception behavior, i.e., the learning network models the risk perception abilities of each particular driver.

B. DRIVING ENVIRONMENT FEATURE ABSTRACTION METHOD

Risk perception is a dynamic interaction between the driver, the vehicle and the driving environment [22], so environmental factors also affect driver risk perception. Therefore, in order to accurately model the risk perception behavior of drivers, we first need to find an effective way to summarize information about the current driving environment and convert that information into a format that can be processed. In this study, traffic-related environmental factors are divided into two types of features, static features and dynamic features, as shown in Table 2. Static features are those environmental factors which remain unchanged for a relative long period of time, such as the type of road, the existence of roadside buildings, traffic infrastructure, and so on. On the other hand, dynamic features are those features of the traffic situation and driver vehicle state which continuously change over time, such as vehicle operation inputs and outputs, the behavior of surrounding vehicles and the presence of obstacles.

TABLE 2. Features of the driving environment.

	Feature class	Feature name
Dynamic features	Ego vehicle	Steering angle
		Brake pedal position
		Gas pedal position
		Longitudinal & lateral acceleration
		steering
	Leading vehicle	Parked state
		Brake lights
		Position
	Obstacles	1_{st} on-coming vehicle
		2_{nd} on-coming vehicle
		Parked vehicles
		Merging vehicles
		Pedestrians & bicycles
Static features	Road structure	Curve
		Uncontrolled intersections
		Controlled intersections

1) STATIC ENVIRONMENTAL FEATURES

Most traffic accidents happen at intersections [29], so road structure should be an important factor affecting the driver's perception of risk. The default road structure in our study is

a straight road with two lanes, one for traffic moving in each direction. Three special road structures are examined; controlled intersections, uncontrolled intersections and curves. Labeling the road structures in the videos is relatively easy, because static environmental features remain unchanged for relatively long periods of time.

2) DYNAMIC ENVIRONMENTAL FEATURES

The state of the driver's vehicle (the ego vehicle) is obtained from the vehicle's on-board diagnostics (OBD). Although the experiment's participants cannot see the speed statistics or the vehicle operation data for each driver, how the vehicle moves in the video affects their risk sensitivity and risk perception. When the ego vehicle is moving at high speed, study participants had a tendency to perceive a higher level of risk. Therefore, the dynamic features include six of the vehicle's operation parameters, representing the movement or operation of the ego vehicle.

The second type of dynamic environmental features are those related to the traffic situation, such as the presence of other vehicles or obstacles. Since drivers assess risk based on the current driving environment, real-time automotive object detection is used to summarize these traffic features in a quantitative form. In this study we adopt a deep learning network-based, real-time object detection method called YOLO [30], [31]. By using YOLO, the locations of the vehicles and obstacles listed in Table 2 can be automatically obtained, as shown in Fig. 4. Compared to other state-of-the-art detection methods, YOLO is an end-to-end detection method, in which the input is an image and the output is all of the detected objects in specified classes, as well as their positions in the image.

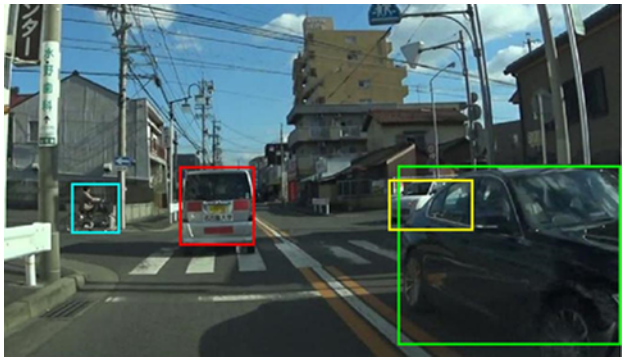


FIGURE 4. Environmental feature abstraction using YOLO.

However, YOLO also collects a lot of extraneous information. For example, it also detects vehicles parked away from the road and pedestrians who are not along the route of the vehicle. These redundant features are generally ignored by drivers and do not affect driver risk perception. As a result, this redundant information is the noise which may hinder the LSTM to learn which features are vital for predicting a driver's risk perception behavior. Therefore, we must manually remove the objects which are unrelated to risk perception to make the data sufficiently clean for the LSTM network. The information which will be input to the LSTM is shown in Fig. 5. The red point is the position of the camera and the green dotted line is the forward-looking view of the driver. The objects in the colored bounding boxes are the objects detected by YOLO which are preserved for risk perception modeling. The green and yellow bounding boxes are the first and second vehicles detected on the opposite side of the road. The red bounding box represents a leading vehicle traveling in the same direction as the ego vehicle. The light blue bounding boxes identify pedestrians, bicycles or motorbikes, which are assumed to have random directions of travel. The dark blue bounding box represents a parked vehicle or motorbike on the driver's side of the road, while the white bounding box identifies a vehicle about to merge into the road that may affect the current operating status of the ego vehicle. The bounding boxes contain traffic information which may be detected visually by the participants when they watch the video clips. By using YOLO, visual traffic information can be transformed into quantitative information which can be processed later by the deep-learning network.

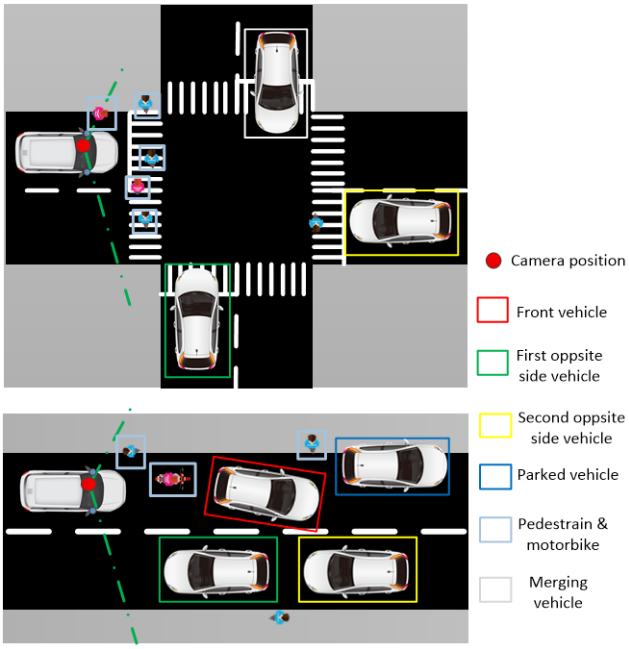


FIGURE 5. Traffic information retained for risk perception modeling.

III. MODELING RISK PERCEPTION BEHAVIOR USING LSTM

A. GROUP AND INDIVIDUAL FEATURES OF THE RISK PERCEPTION DATA

Fig. 6 shows the risk perception distributions for each of the four participant groups. We can see that when viewing the same driving videos, the driving instructors were the most sensitive to risk, followed by the experienced drivers, and then the elderly drivers. The novice drivers were much less sensitive to the risk compared to the other three groups.

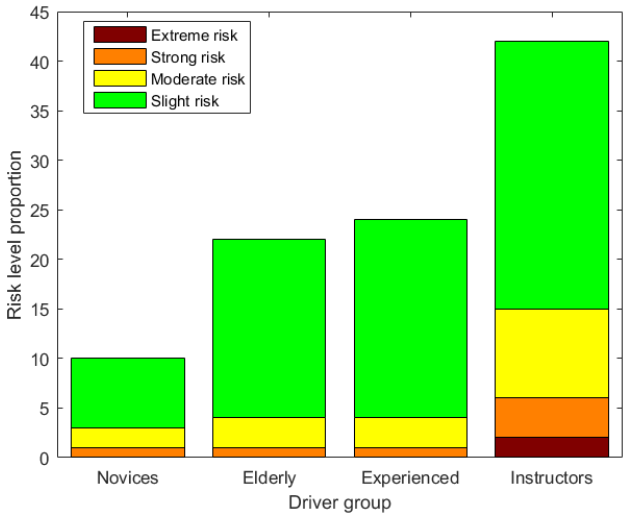


FIGURE 6. Statistical distributions of risk perception for each driver group.

However, when examining the details of each individual driver's risk perception data, we can see that there is considerable variation in risk perception ability within each group. In Fig. 7, we compare the statistical properties of the risk perception of each group when viewing the same video. The vertical axis of Fig. 7 represents the ratios of perceived risk for each group. The ratios of perceived risk represent how many studies participants treat current time point as risk, 0 represent for no one perceived current time as risk and 1 represents all participants perceive the risk. Since risk perception among the groups differs, and it also differs even within the same group, we can see that risk perception is not consistent, even among drivers with similar profiles. Therefore, in this paper we use each participant's risk perception data to construct individualized risk perception models. Before training the learning network, we converted the raw risk perception data from five levels of perceived risk into a binary form (perceived risk or no perceived risk).

B. RISK PERCEPTION MODELING USING LSTM

Risk perception by drivers is based on previous and current traffic conditions, as well as on the movement of the driver's vehicle. Just as reading and writing are the cognitive result of stimulation by the outside environment or reflection on objective stimulation, the behavior of risk perception can be treated as a time sequence information process or, in other words, a long-term dependency. Long Short Term Memory, first proposed by Hochreiter and Schmidhuber [24], is a very effective method of time sequence modeling which can be used to solve long-term dependency problems [25] and to effectively model many human-like behaviors [26]–[28]. The risk perception process is analogous to human linguistic behavior. Just as a human being says or writes a word based on the previous several words, drivers perform risk assessment based on previous traffic situations. For every period of data collection, the collected environmental and vehicle behavior data can be seen as a word generator. Just as human beings cannot understand the overall meaning of a sentence from a single word, but instead require a sequence of words, one is also unable to make a risk judgment based on a single snapshot of traffic information or vehicle state data. Therefore, we utilize 17 kinds of traffic environment feature data over a period of 3 continuous seconds to construct a sequence of 'words', and each 3 second data segment is then defined as 'risky' or 'not risky' by the experiment's participants. The LSTM is designed so that the network will provide the same risk assessment as the participant when encountering a similar combination of 'words'. The use of 3 seconds of feature input was an a priori choice in this study. The longer the time sequence of the input data, the more difficult it is for the network to process the larger number of features, and the harder it is for the network to converge. Additionally, more training data is needed if there is more input data in order to avoid potentially under-fitting the network. If a shorter time sequence of input data is used, the network might not be able to obtain sufficient information for modeling driver risk perception behavior and this could result in network over-fitting.

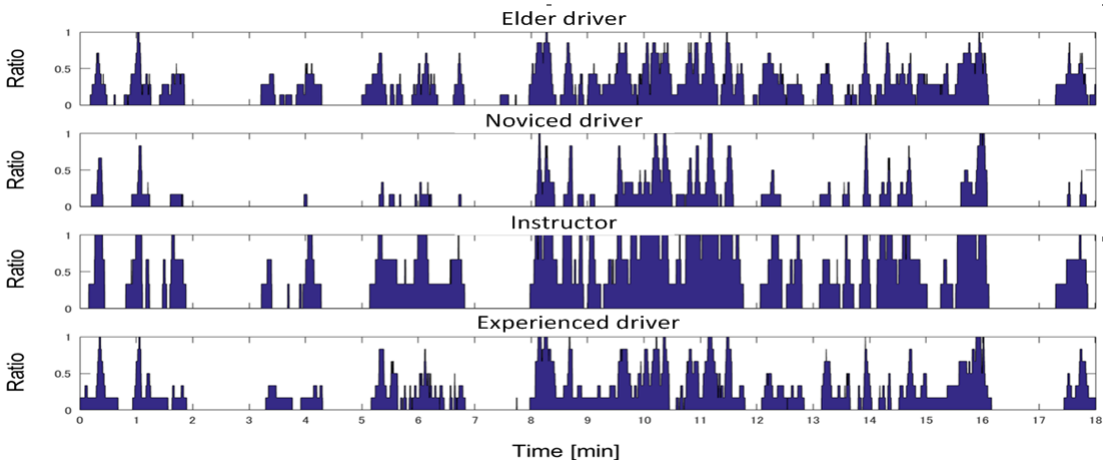


FIGURE 7. Statistical properties of the risk perception data for the members of each group when viewing the same video (over time).

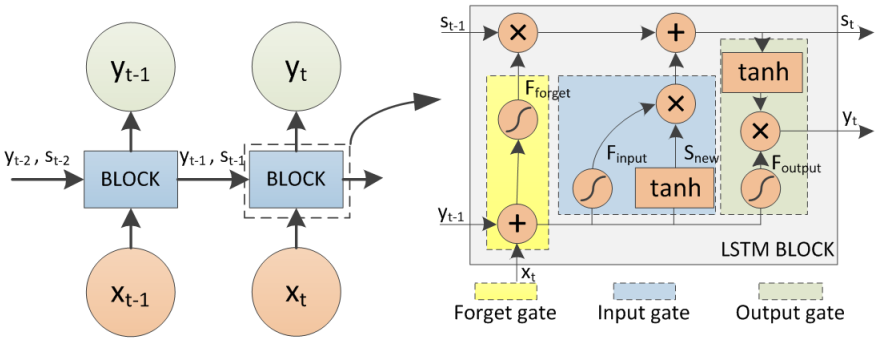


FIGURE 8. LSTM network structure.

In Fig. 8, the framework of the LSTM network and the inner structure of the hidden nodes are illustrated. Although the network structure of LSTM is very similar to an RNN, the major contribution of the LSTM is the introduction of a 'forget gate' cell.

The yellow part of the LSTM block in Fig. 8 is the forget gate cell. The mathematical expression of the forget gate is shown in (1) and (2), where σ represents the standard logistic sigmoid function of the nonlinear block in the forget gate, W_{forget}, b_{forget} are the weight matrix and bias of the forget gate cell respectively, s_{t-1} is the network inner state of the previous time period, and x_t represents the current input to the network. The forget gate cell determines the effect of past information on the current time period.

$$F_{forget} = \sigma(W_{forget} \cdot [s_{t-1}, x_t] + b_{forget}) \quad (1)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

The second part of the LSTM block is the input gate cell, which decides what new information should be added to the new network inner state. Equations (3) and (4) are mathematical representations of the input gate.

$$F_{input} = \sigma(W_{input} \cdot [y_{t-1}, x_t] + b_{input}) \quad (3)$$

$$S_{new} = tanh(W_{new} \cdot [h_{t-1}, x_t] + b_{new}) \quad (4)$$

The network's inner state can be updated by the output of the forget gate and input gate, as defined in (5), where $\otimes$ represents the scalar product of two vectors.

$$s_t = F_{forget} \otimes s_{t-1} + F_{input} \otimes S_{new} \quad (5)$$

The last component of the LSTM block is the output gate cell as defined in (6). The output gate decides which parts of the network state and current information to output.

$$y_t = \sigma(W_{output} \cdot [y_{t-1}, x_t] + b_{output}) \otimes tanh(s_t) \quad (6)$$

In each time step, input x_t will be evaluated by the LSTM block, which will produce the output y_t based on the current input and the previous network inner state. In this study, network input x_t contains the information listed in Table 2. All of the parameters are normalized into 0 to 1 before being input into the network. Three seconds of risk perception data (30 data points) will be packed into one packet of data and used as the LSTM network's input. The output of the LSTM is a predicted perceived risk. During each training step, the network will compare its output with the participant's risk perception data and the calculated deviation will be used to adjust the inner parameters (W_{forget}, b_{forget}, W_{input}, b_{input}, W_{new}, b_{new}, W_{output}, b_{output}) of the LSTM block. This process is known as 'back-propagation through time' (BPTT) [32]. After BPTT, the LSTM will output the risk prediction results, which performs the similar risk perception behavior of the participants'.

When training the LSTM network, a 'sliding window' is used to construct the input data, as shown in Fig. 9. The sliding window contains 30 data points and each sliding step consists of 4 data points. By using a sliding window, the packed input data is in a time sequence. For each time step, the packet data will be updated with new data for the next 0.4 seconds. This sliding window enables the network to learn the driver's risk perception behavior every 0.4 seconds based on the previous 3 seconds of environmental features. This is consistent with the human visual reaction time, which ranges from about 0.3 seconds to 0.8 seconds [33].

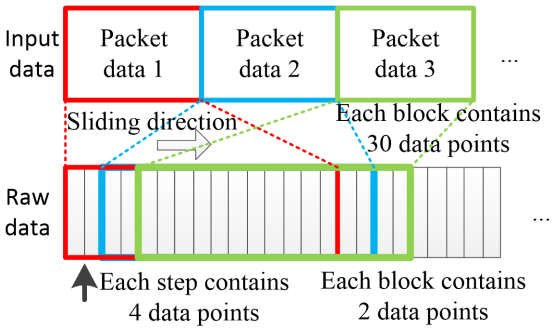


FIGURE 9. Sliding window mechanism for training data composition.

The components of the LSTM network are constructed using TensorFlow [34], which is an open-source, deep-learning calculation platform, and the LSTM block is constructed using the LSTM cell provided by TensorFlow, along with the network hyper-parameters shown in Table 3.

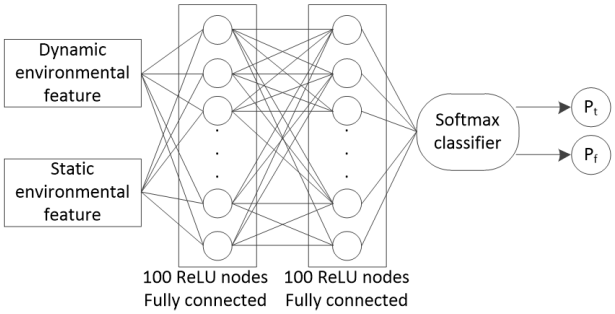


FIGURE 10. Neural network for risk perception modeling.

TABLE 3. Hyper-parameters of the LSTM.

Parameter name	Parameter value
# of nodes in each hidden layer	50,75,100,128
Initial forget bias	1
# of output classes	2
Learning rate	0.0001
Optimizer	Gradient Descent
Loss	Cross entropy
Output layer	Softmax

The final output of the LSTM will be put into a Softmax binary classifier to calculate the probability of risk. The detail modeling results will be discussed in the next section.

IV. RESULTS

The risk perception modeling results by using different methods are detailed in this section. First, the modeling performance of a neural network (NN), a support vector machine (SVM) [35] and two types of LSTM network are compared. Second, based on the results of this comparison, a discrete coding method for the LSTM is proposed in order to improve modeling accuracy. Lastly, the individual modeling results and risk contributing factors are given.

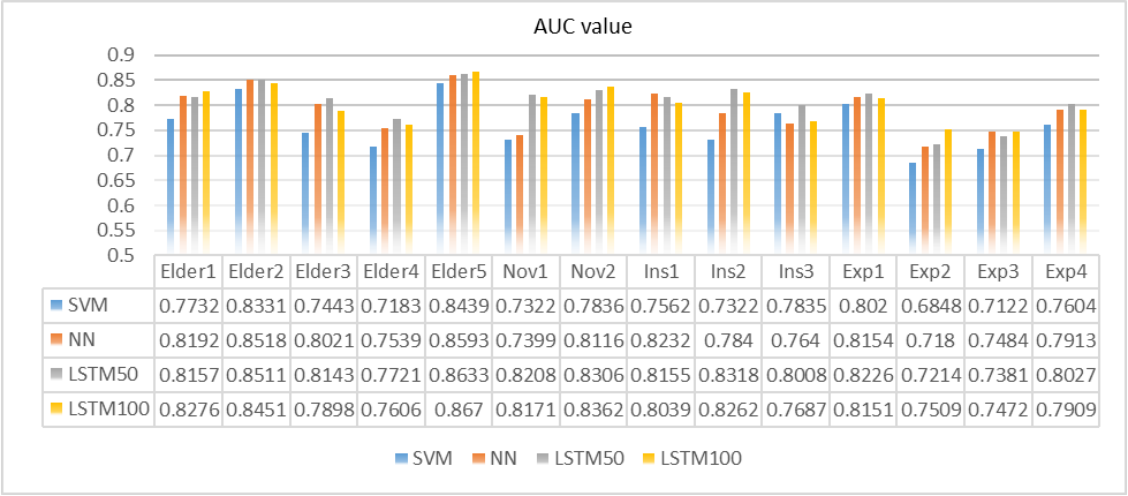
A. TRAINING DATA

As the risk perception model is approximating the risk perception of an individual driver, the training and testing data consist of each participant's own risk perception assessment data. We divided each participant's data into six groups (four training groups, one validation group and one testing group). Each group contained about 10,000 data samples. The training process includes cross-validation, so each data group will be treated as validation or testing data group to improve the generalization ability and robustness of the network.

B. COMPARISON OF DIFFERENT MODELING METHOD

Since the type of learning methods affects the modeling results, in this section we compare four kinds of learning structures; a multi-layer neural network (NN), two LSTMs (each with the different number of hidden nodes), and the kernel-based method-support vector machine (SVM). The structure of the NN is shown in Fig. 10. Two Rectified Linear Unit (ReLU) layers [36] compose the hidden layers of the NN, and the output layer is a binary classifier with Softmax. P_t and P_f represent risk prediction probabilities for the next point in time. P_t shows the probability given by the learning network that the participant will detect or feel risk during the next time period, while P_f is the probability that the participant will not detect any risk. Our SVM was constructed using the MATLAB SVM toolbox [37]. The output layer of the LSTMs and the SVM is also a binary classifier using Softmax, so the output of these two methods is also in the form of risk prediction probabilities P_t and P_f.

14 of the participants (five elderly drivers, two novices, three driving instructors and four experienced drivers) was selected as the analysis sample. As the modeling structure proposed in this paper is in fact a type of classifier, which simulates human behavior when classifying the current situation as either risky or not risky, the area under the curve of receiver operating characteristics (AUC) [38] and the receiver



	Elder1	Elder2	Elder3	Elder4	Elder5	Nov1	Nov2	Ins1	Ins2	Ins3	Exp1	Exp2	Exp3	Exp4
SVM	0.7732	0.8331	0.7443	0.7183	0.8439	0.7322	0.7836	0.7562	0.7322	0.7835	0.802	0.6848	0.7122	0.7604
NN	0.8192	0.8518	0.8021	0.7539	0.8593	0.7399	0.8116	0.8232	0.784	0.764	0.8154	0.718	0.7484	0.7913
LSTM50	0.8157	0.8511	0.8143	0.7721	0.8633	0.8208	0.8306	0.8155	0.8318	0.8008	0.8226	0.7214	0.7381	0.8027
LSTM100	0.8276	0.8451	0.7898	0.7606	0.867	0.8171	0.8362	0.8039	0.8262	0.7687	0.8151	0.7509	0.7472	0.7909

FIGURE 11. AUC values for each method when modeling each driver.

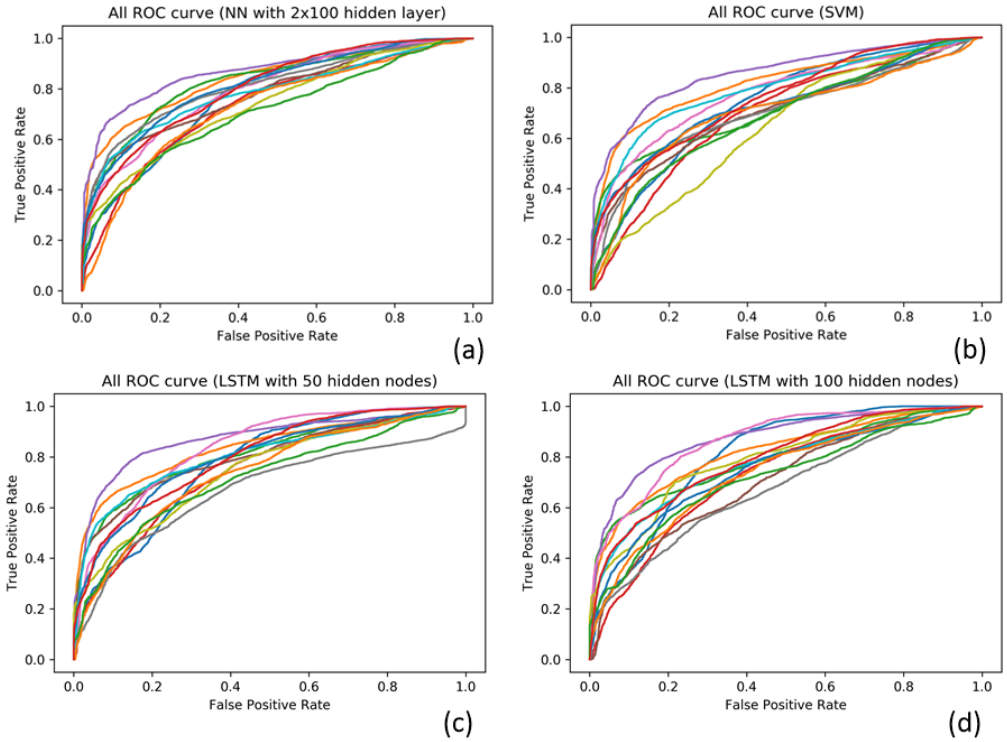


FIGURE 12. ROC for each modeling method when modeling each driver.

operating characteristics curve (ROC) [39] are adopted as the evaluation indices. The modeling performance of each method is compared in Table 4. The appropriateness of using the criterion of the AUC to judge whether or not a classifier is effective depends on the context of the analysis. For example, in medical diagnosis a very high AUC is defined as 0.95 or even higher. However, in the fields of applied psychology and the prediction of future behavior, an AUC value above 0.7 is considered to be very effective [40]. So in this paper, we suggest that an AUC above 0.7 certifies that a classification method can model risk perception well. As we can see in Table 4, all of the methods we evaluated were effective for modeling the risk perception of the various groups of drivers, but the proposed LSTM based methods achieved the best performance.

TABLE 4. AUCs for each modeling method.

Learning network category	NN	SVM	LSTM with 50 hidden nodes	LSTM with 100 hidden nodes
All participants	0.792	0.761	**0.807**	0.803
Elderly	0.783	0.817	**0.823**	0.818
Novices	0.776	0.758	0.825	**0.826**
Instructors	0.790	0.757	**0.816**	0.799
Experienced	0.768	0.740	0.771	**0.776**

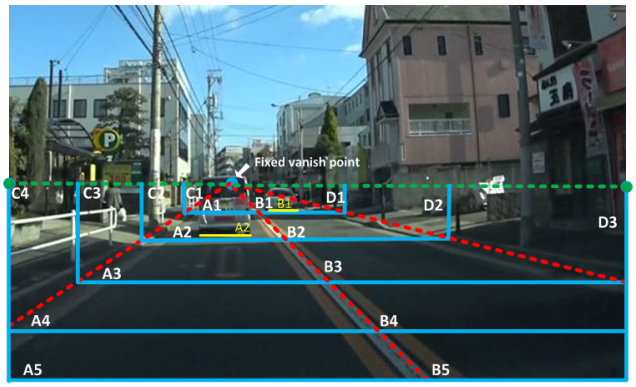


FIGURE 13. Dividing the driving environment into districts.

The AUC values and ROC curves using each learning method, for each of the participants in the experiment, are shown in Fig. 11 and Fig. 12, respectively. From these results, we can see that each method achieves different modeling performance for each individual. Overall, the LSTM with 50 hidden nodes and the LSTM with 100 hidden nodes achieved the best classification performance.

C. IMPROVING MODELING PERFORMANCE

Theoretically, LSTM should be more effective when modeling this type of time series information than the other methods. However, based on our results reported above, we found that for some drivers the LSTM approach was less effec-

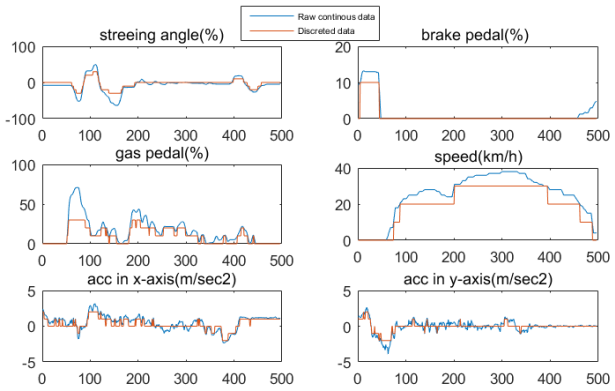


FIGURE 14. Samples of discrete vehicle operation data.

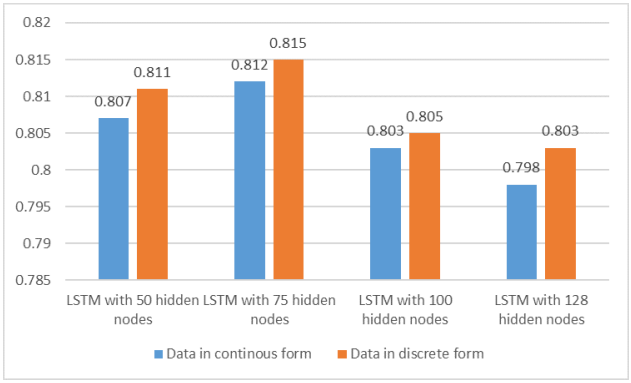


FIGURE 15. Comparison of risk assessment modeling performance using continuous and discrete data, for LSTMs with different numbers of hidden nodes.

tive than the NN method. To investigate this phenomenon we performed another experiment, in which we supplied the training data in binary form to train the LSTM, using 0 or 1 to represent traffic environment factors (no present vs. present) instead the exact positions of the objects. As a result, we found that much higher AUC values were achieved. We hypothesize that due to current size of the training data, the raw training data in continuous form may lower the LSTM's modeling performance. As a result, the positions of the surrounding traffic objects need to be encoded in a more abstract way, instead of directly inputting raw position data into the LSTM. Inspired by the improved results when using binary traffic environment information, we tried replacing the exact positions of the surrounding objects with general locations. As shown in Fig. 13, the original positions of the traffic objects (the yellow lines) were utilized to deduce their location district. Then each frame of the assessment videos was divided into districts and the position supplied by YOLO was used to deduce which district the traffic objects belong to.

The red dotted lines are the detected lane boundaries and the green dotted line is the horizon, which intersects the vanishing point. The Fig. 13 is now divided into four districts (A,B,C and D). The A district is the lane used by the ego vehicle, B is the opposite lane, and C and D are the areas between the green line and red line adjacent to the driver's lane and the opposite lane, respectively. Traffic objects in the video frame were then coded using the district in which the yellow line under the objects was located, as shown in Fig. 13. When using this improved traffic coding method, traffic factor information was reduced from 4 to 2 dimensions, and each dimension's variability was also limited. Information about the state of the ego vehicle was also made discrete,

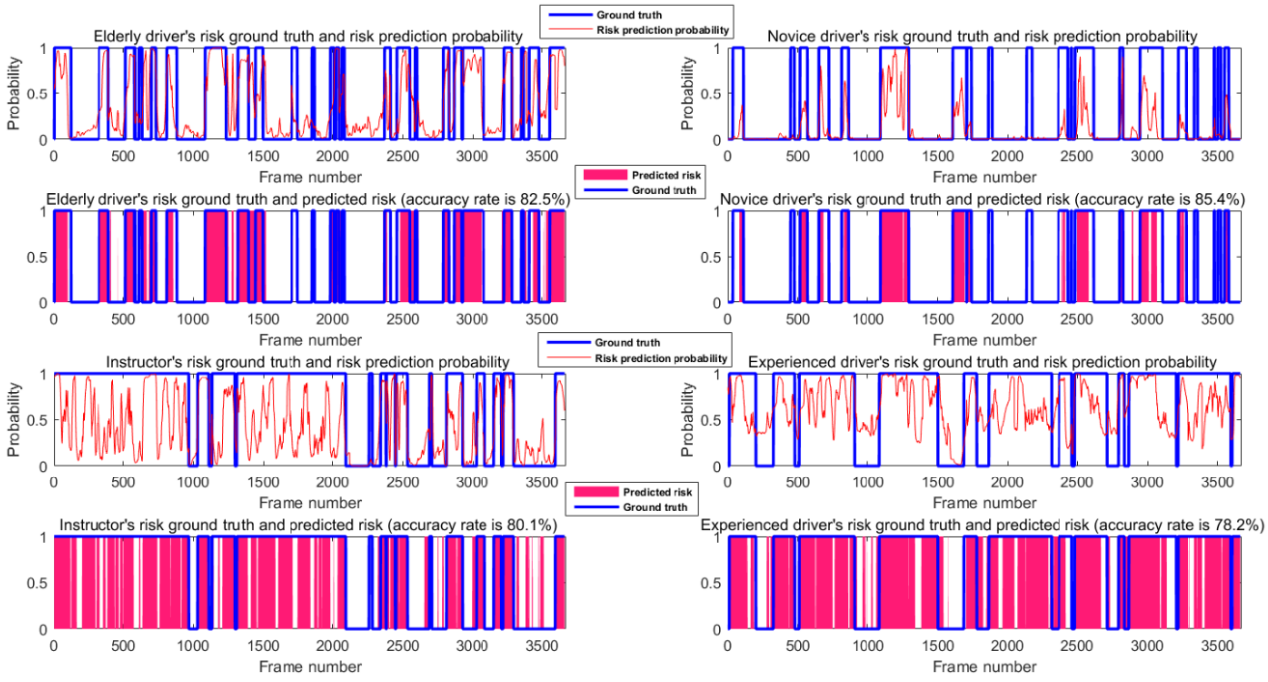


FIGURE 16. Individual risk modeling results.

FIGURE 17. Risk prediction data correlated to images of traffic conditions.

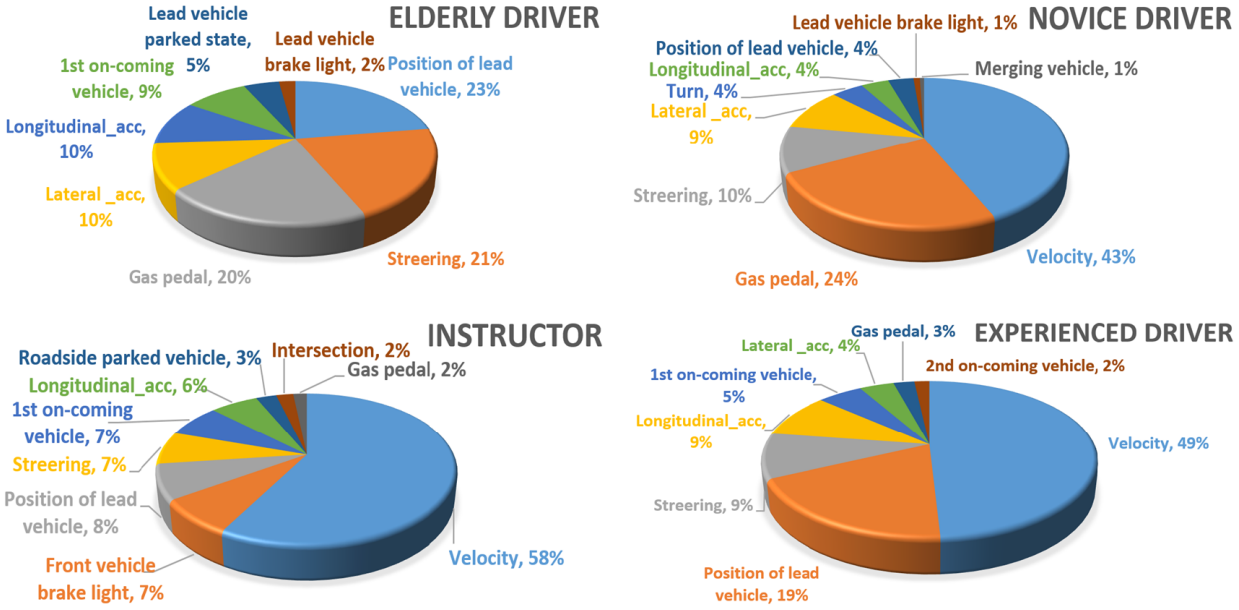


FIGURE 18. Importance of various environmental factors on the risk assessments of participants from each of the four types of drivers.

based on distributions of the vehicle operation information. An example of the discrete data is shown in Fig. 14.

A comparison of the results when using detailed position and operation data versus district location and discrete vehicle operation data is shown in Fig. 15. As we can see, the performance of each of the LSTM methods was improved by using less precise data. We also investigated the effect on modeling performance of using LSTMs with different numbers of hidden nodes. We found that an LSTM with 75 hidden nodes achieved the best performance.

D. INDIVIDUALIZED RISK PERCEPTION PREDICTION RESULTS

Modeling results for selected four individuals are shown in Fig. 16. In order to illustrate learning network performance, 3,600 sequential data samples were selected. The blue line represents the ground truth, based on evaluations result of the same video by the selected four participants. The red line represents the raw risk prediction probability of the LSTM for each individual. As mentioned before, the output layer of the LSTM is a binary Softmax classifier. Therefore, the raw result of the LSTM is a risk probability. A threshold of the probability is needed to determine whether the result represents a true or false risk prediction. The optimal threshold can be obtained by maximizing the AUC. The red area is the prediction result for each driver, which is calculated by using the optimal threshold. Although the risk prediction ability of each driver is different, our deep learning-based risk modeling method can model each driver's risk perception well, with an average accuracy rate of 81.55%. Examples of detailed risk prediction data as correlated to images of the observed traffic conditions are shown in Fig. 17 and more samples are given in Appendix.A and Youtube.

E. INDIVIDUALIZED RISK PERCEPTION FACTOR CONTRIBUTION ANALYSIS

The participants in the experiment made their risk assessments based on several environmental factors. In our study we also analyzed which factors were the most important for each individual when making their risk assessments, using our

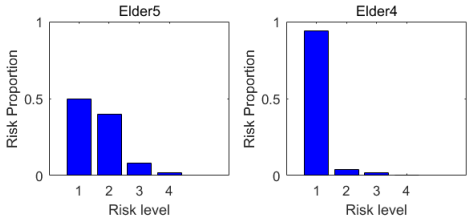


FIGURE 19. Risk assessment distributions of two participants (1 = the lowest level of perceived risk).

proposed model. For each participant's set of risk estimation data, we eliminated specific environment factors to identify their effect on the final risk perception results. The effect score was then calculated using (7).

$$F_n = \frac{|P_n - P_o|}{\sum_{j=1}^{N}(P_j - P_o)} \cdot sigmoid(P_o) \tag{7}$$

F_n represents the contribution or weight of the n^{th} environmental factor on a participant's risk assessment in each risk prediction step. For example F_1 is the contribution of the 1^{st} environmental factor, which is steering angle. P_n is the risk perception probability when the n^{th} environmental factor is eliminated from the test data, P_o is the original risk perception probability calculated using the test data containing all of the environmental factors. N is the number of environmental factors (in this study there were 17 such factors). The sigmoid is a nonlinear function related to P_o (this sigmoid function was introduced to calculate the degree to which a particular contribution factor effected a participant's risk assessment). After calculating, the environmental factors with biggest contribution will be treat as the most important factors for the individual to make risk perception. We selected representatives of each of the four types of drivers and calculated the weight of each environmental factor on their risk assessments. Distributions of the influence of the various environmental factors are shown in Fig. 18.

V. DISCUSSION

In this section, we want to provide a more detailed review of our modeling results and provide additional perspectives.

A. MODEL PERFORMANCE

As shown in the Fig. 15, the best AUC for the enhanced version of the proposed model was 0.815, which demonstrates that our modeling method can effectively classify the risk perception behavior of drivers. Despite these favorable results, some issues remain which need to be addressed. First, our enhanced modeling method is in fact a binary classifier, so it cannot provide complex information about risk perception such as the level of risk perceived or the risk category directly. Second, when training the models we also discovered that the modeling results were influenced by the risk assessment performance of the participants as they viewed the videos. As shown in Fig. 19, we picked two participants with

FIGURE 20. Samples of true risk perception correlated to images of traffic conditions(Upper left: risk predicted in intersection; Upper right: risk from the second opposite vehicle; Bottom left: risk from the roadside pedestrian; Bottom right: risk from the front parked vehicle).

FIGURE 21. Samples of true risk perception correlated to images of traffic conditions(Upper left; risk from the parked vehicle. Upper right; risk from the merging vehicle. Bottom left; risk from the crossing intersection. Bottom right; risk from the parked vehicle).

FIGURE 22. Samples of true risk perception correlated to images of traffic conditions (Upper left; risk from the left turn. Upper right; risk from the left turn. Bottom left; risk from the pedestrian.Bottom right; risk from the pedestrian).

obviously different risk assessment distributions when viewing the same 44 videos. Clearly, participant Elder4 perceived that most of the driving situations he observed were safe. When we modeled the risk assessment behavior of these two participants, the AUC of the model was 0.867 for Elder5 and 0.76 for Elder4. The likely explanation for this difference in

FIGURE 23. Samples of true risk perception correlated to images of traffic conditions (Upper left; risk from the stop of the leading vehicle. Upper right; risk from the stop of the leading vehicle. Bottom left; risk from the right turn.Bottom right; risk from the right turn).

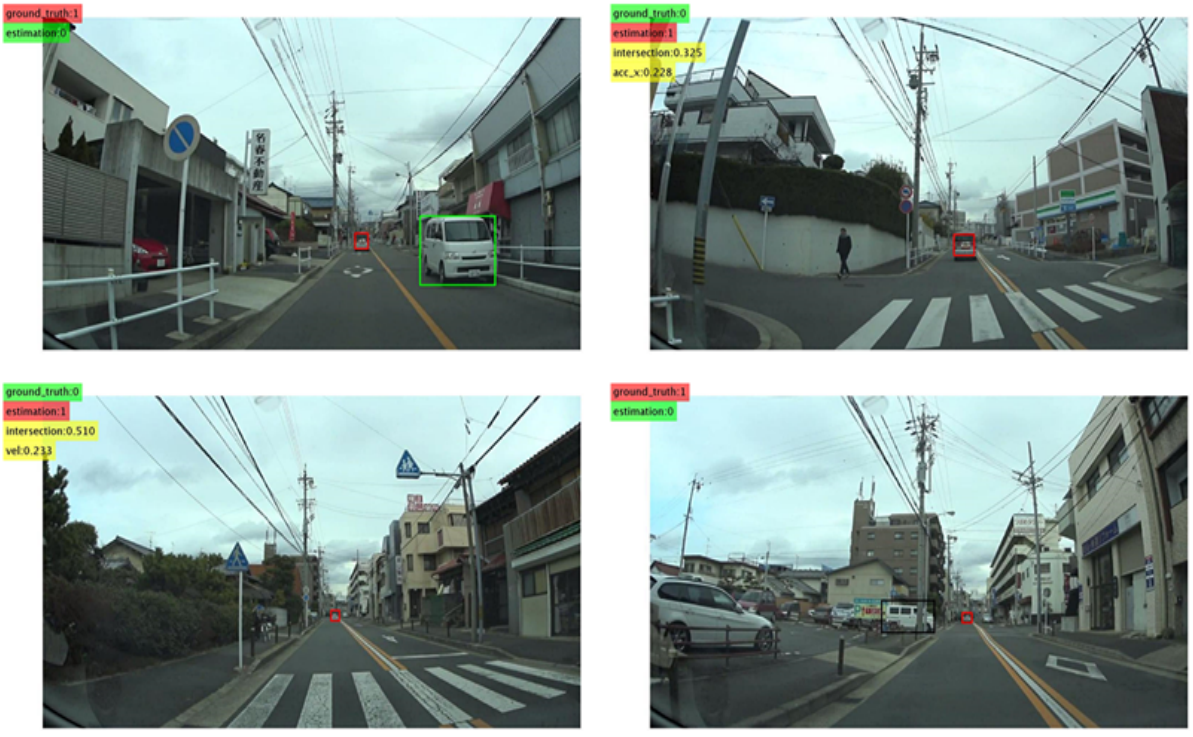


FIGURE 24. Samples of false risk perception correlated to images of traffic conditions (Upper left; risk perception model did not predict any risk. Upper right; risk perception model predict the risk approaching the intersection while the study participant did not. Bottom left; same as the upper right figure.Bottom right; risk perception model did not predict the risk but after 1 seconds later).

performance is that when observers underestimate potential risk, the result is that the training data contains insufficient risk features for the model to learn from. As a result, the LSTM will have difficulty understanding which features cause a particular participant to perceive the presence of risk.

B. LSTM NETWORK STRUCTURE'S EFFECT

In the course of this study, we changed the number of hidden nodes of the LSTM and the form of the input training data in order to find the optimal LSTM for our modeling task. As the current optimal AUC is 0.815 (using common LSTM structure [41] with 75 hidden nodes), there is much room for improvement in our deep learning-based risk perception modeling method. There are many ways to configure the network structure of an LSTM; there are bi-directional LSTMs [42], stack bi-directional LSTMs [43], LSTMs with attention [26], and so on. Comparing the performance of these various LSTM architectures will require the consideration of many factors such as the recomposition of the training data and design of the attention function. These investigations will be the focus of our future research.

C. RESULTS OF INDIVIDUALIZED RISK MODELING ANALYSIS

In this paper we constructed a risk perception model and tested it using drivers from four different driver categories (novices, elderly drivers, experienced drivers and driving instructors). We then analyzed which environmental factors most strongly influenced risk perception among representatives of each driver category. As we can see in Fig. 18, the elderly driver was most focused on the leading vehicle, followed by the two ego vehicle states of steering and gas pedal operation, which is consistent with the results of a study by Siren and Kjær which found that older drivers tended to think of risk as something external [6]. We can also see in Fig. 6 that the novice drivers were much less sensitive to risk compared to the drivers from the other three groups. We can also observe Fig. 18 that the selected novice driver was most sensitive to velocity of the ego vehicle, possibly because novice drivers feel anxious when driving at high speed. We can also see that risk factors from the outside environment made up only 10% of the important risk factors for our novice driver, leading to the conclusion that this driver has little ability to perceive or predict risk in a dynamic traffic environment. The driving instructor and experienced driver were also very sensitive to speed, but the percentages of important risk factors originating from the outside environment were 27% and 26%, respectively, for these two drivers.

VI. CONCLUSION

In this paper, we proposed a method of modeling driver risk perception based on a deep learning network. We believe that this study has made several contributions to the fields of risk perception and driver behavior modeling: As far as we know, this is the first time a deep learning network has been used for risk perception modeling in a study which included factors related to both the status of the ego vehicle and information about the external driving environment. Our research integrates the driving environment in order to construct a risk perception model which can break down the factors influencing risk perception to the micro-factor level. As a result, our proposed modeling method allows us to recognize situations in which a particular driver will feel they are at risk, or will perceive the presence of risk. We were also able to realize end-to-end risk perception analysis for our model by using YOLO, which is a real time, environmental factor abstracting method, to detect the presence of environmental features along the route and establish their locations. The inputs to our model are video data from the driver's view of the road ahead and data about the current status of the driver's own vehicle. Using a deep learning based network, the output, which is binary data representing the presence or absence of perceived risk, can be used to predict a risk perception result similar to the driver's. The real-time capability of our model means it could be used to provide a human behavior-based foundation for safe decision making for future intelligent driving assistance systems.

On the other hand, our research also contains some limitations. The selected environmental factor categories are not sufficient to model the entirety of the driving environment. In this study, we only collected the data needed to model driver behavior on a selected, relatively uniform experimental road. It will be difficult to adapt our current model to other driving environments, such as multi-lane highways, country roads (without clear road boundary), night driving, and so on. What's more, the results of our analysis are relatively simple, because we only wanted to determine whether or not the driver will perceive risk. The limited size and depth of the training data also make it difficult for the LSTM network to model more complex features. As a result, when using the currently available data set, it is difficult to train the network to obtain the desired risk category details. In the future, we need to collect environmental data from more complex driving situations in order to make our model more general. The dimensions of the training data also need to be expanded to satisfy the need for complex feature abstraction.

APPENDIX
SAMPLES OF RISK PERCEPTION CORRELATED TO IMAGES OF TRAFFIC CONDITIONS

A. SAMPLES OF TRUE AND FALSE RISK PERCEPTION CORRELATED TO IMAGES OF TRAFFIC CONDITIONS

Ground truth represent the risk perception made by the study participants, the proposed modeling makes estimation. The yellow bounding box are the top two contributor for the risk perception. A few representative samples are listed in the Appendix.A. We also offer more examples on the Youtube. The website address is https://www.youtube.com/watch?v=F8BhHtFxS18.

ACKNOWLEDGMENT

This work was supported in part by a Grant-in-Aid for Scientific Research (C) under Grant 15K002231 from the Japan Society for the Promotion of Science (JSPS), in part by the National Natural Science Foundation of China (Grant 61300101), and in part by the Key Research Plan of Jiangsu Province (BE2017035). The authors would like to thank the Vehicle Engineering Development Division of Mitsubishi Motors for their valuable research assistance.

REFERENCES

[1] World Health Organization, "Global status report on road safety 2015," *World Heal. Org.*, vol. 19, no. 2, p. 150, 2015.

[2] N. Morisset, F. Terrade, and A. Somat, "Perceived self-efficacy and risky driving behaviors: The mediating role of subjective risk judgment," *Swiss J. Psychol.*, vol. 69, no. 4, pp. 233–238, 2010.

[3] L. Eboli, G. Mazzulla, and G. Pungillo, "How to define the accident risk level of car drivers by combining objective and subjective measures of driving style," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 49, pp. 29–38, Aug. 2017.

[4] M. S. Horswill and F. P. McKenna, "Drivers' hazard perception ability: situation awareness on the road," in *A cognitive approach to situation awareness: theory and application*, S. Tremblay and S. Banbury, Eds. Aldershot, U.K.: Ashgate, 2004, pp. 155–175. [Online]. Available: http://centaur.reading.ac.uk/14289/

[5] L. Sjoberg, B. Moen, and T. Rundmo. (2004). *Explaining Risk Perception*. An Evaluation of the Psychometric Paradigm in Risk Perception Research. [Online]. Available: http://www.svt.ntnu.no/psy/torbjorn.rundmo/psychometric_paradigm.pdf

[6] A. Siren and M. R. Kjær, "How is the older road users' perception of risk constructed?" *Transp. Res. F, Traffic Psychol. Behav.*, vol. 14, no. 3, pp. 222–228, 2011.

[7] K. Nevelsteen, T. Steenberghen, A. Van Rompaey, and L. Uyttersprot, "Controlling factors of the parental safety perception on children's travel mode choice," *Accident Anal. Prev.*, vol. 45, pp. 39–49, Mar. 2012.

[8] V. Rafaely, J. Meyer, I. Zilberman-Sandler, and S. Viener, "Perception of traffic risks for older and younger adults," *Accident Anal. Prev.*, vol. 38, no. 6, pp. 1231–1236, 2006.

[9] C. L. Wright and K. Silberman, "Media influence on perception of driving risk and behaviors of adolescents and emerging adults," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 54, pp. 290–298, Apr. 2018.

[10] J. L. Machado-León, J. De Oña, R. De Oña, L. Eboli, and G. Mazzulla, "Socio-economic and driving experience factors affecting drivers' perceptions of traffic crash risk," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 37, pp. 41–51, Feb. 2016.

[11] I. O. Lund and T. Rundmo, "Cross-cultural comparisons of traffic safety, risk perception, attitudes and behaviour," *Safety Sci.*, vol. 47, no. 4, pp. 547–553, 2009.

[12] D. Crundall *et al.*, "Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard," *Accident Anal. Prev.*, vol. 45, pp. 600–609, Mar. 2012.

[13] V. Dixit, G. W. Harrison, and E. E. Rutström, "Estimating the subjective risks of driving simulator accidents," *Accident Anal. Prev.*, vol. 62, pp. 63–78, Jan. 2014.

[14] R. Takahashi, M. Kobayashi, T. Sasaki, Y. Yokokawa, H. Momose, and T. Ohhashi, "Driving simulation test for evaluating hazard perception: Elderly driver response characteristics," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 49, pp. 257–270, Aug. 2017.

[15] D. Kaber, S. Jin, M. Zahabi, and C. Pankok, "The effect of driver cognitive abilities and distractions on situation awareness and performance under hazard conditions," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 42, pp. 177–194, Oct. 2016.

[16] G. Underwood, D. Crundall, and P. Chapman, "Driving simulator validation with hazard perception," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 14, no. 6, pp. 435–446, 2011.

[17] A. Borowsky, D. Shinar, and T. Oron-Gilad, "Age, skill, and hazard perception in driving," *Accident Anal. Prev.*, vol. 42, no. 4, pp. 1240–1249, 2010.

[18] M. R. Endsley, "Situation awareness and aging," in *A Cognitive Approachto Situation Awareness: Theory and Application*, S. Banbury and S. Tremblay, Eds. Aldershot, U.K.: Ashgate, 2004, pp. 317–341.

[19] G. Underwood, N. Phelps, C. Wright, E. Van Loon, and A. Galpin, "Eye fixation scan paths of younger and older drivers in a hazard perception task," *Ophthalmic Physiol. Opt.*, vol. 25, no. 4, pp. 346–356, 2005.

[20] D. Crundall, "Hazard prediction discriminates between novice and experienced drivers," *Accident Anal. Prev.*, vol. 86, pp. 47–58, Jan. 2016.

[21] A. Gugliotta *et al.*, "Are situation awareness and decision-making in driving totally conscious processes? Results of a hazard prediction task," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 44, pp. 168–179, Jan. 2017.

[22] M. Liu, Y. Chen, G. Lu, and Y. Wang, "Modeling crossing behavior of drivers at unsignalized intersections with consideration of risk perception," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 45, pp. 14–26, 2017.

[23] X. Zhao, Q. Li, D. Xie, J. Bi, R. Lu, and C. Li, "Risk perception and the warning strategy based on microscopic driving state," *Accident Anal. Prev.*, vol. 118, pp. 154–165, Sep. 2018.

[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[25] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 1310–1318.

[26] K. Xu *et al.*, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2048–2057.

[27] D. Bahdanau, K. Cho, and Y. Bengio. (2014). "Neural machine translation by jointly learning to align and translate." [Online]. Available: https://arxiv.org/abs/1409.0473

[28] J. Morton, T. A. Wheeler, and M. J. Kochenderfer, "Analysis of recurrent neural networks for probabilistic modeling of driver behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1289–1298, May 2017.

[29] *Traffic Safety Facts: 2015*, Nat. Highway Traffic Safety Admin., Washington, DC, USA, 2015.

[30] J. Redmon and A. Farhadi. (2018). "YOLOv3: An incremental improvement." [Online]. Available: https://arxiv.org/abs/1804.02767

[31] J. Redmon and A. Farhadi. (2016). "YOLO9000: Better, faster, stronger." [Online]. Available: https://arxiv.org/abs/1612.08242

[32] M. C. Mozer, "A focused backpropagation algorithm for temporal pattern recognition," *Complex Syst.*, vol. 3, no. 4, pp. 349–381, 1989.

[33] A. Mirzaei, S. M. Khaligh Razavi, M. Ghodrati, S. Zabbah, and R. Ebrahimpour, "Predicting the human reaction time based on natural image statistics in a rapid categorization task," *Vis. Res.*, vol. 81, pp. 36–44, Apr. 2013.

[34] G. Brain. (2018). *TensorFlow*. [Online]. Available: https://www.tensorflow.org/

[35] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[36] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," *J. Mach. Learn. Res.*, vol. 15, no. 4, pp. 315–323, 2011.

[37] Mathworks. (2018). *Support Vector Machines for Binary Classification*. [Online]. Available: https://ww2.mathworks.cn/help/stats/support-vector-machines-for-binary-classification

[38] C. X. Ling, J. Huang, and H. Zhang, "AUC: A better measure than accuracy in comparing learning algorithms," in *Proc. 16th Conf. Can. Soc. Comput. Stud. Intell., AI* (Lecture Notes in Computer Science), vol. 2671, Halifax, Canada, Jun. 2003, pp. 329–341. [Online]. Available: https://www.researchgate.net/publication/221442229_AUC_A_Better_Measure_than_Accuracy_in_Comparing_Learning_Algorithms

[39] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

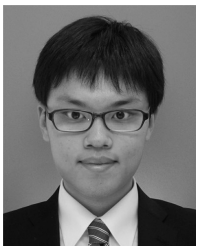
[40] M. E. Rice and G. T. Harris, "Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r," *Law Hum. Behav.*, vol. 29, no. 5, pp. 615–620, 2005.

[41] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proc. Interspeech*, Sep. 2014, pp. 338–342. [Online]. Available: https://www.researchgate.net/publication/279714069_Long_short-term_memory_recurrent_neural_network_architectures_for_large_scale_acoustic_modeling

[42] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.

[43] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. ICASSP*, vol. 3, 2013, pp. 6645–6649.

PENG PING received the B.S. degree in automation from the Beijing University of Chemical Technology, Beijing, China, in 2010, and the M.S. degree in automation from the Nanjing University of Science and Technology, Nanjing, China, in 2013. He is currently pursuing the Ph.D. degree with Southeast University, Nanjing, China. In 2017, he went to Nagoya University as a joint Ph.D. Student. From 2013 to 2015, he was a Research and Development Engineer as part of the Cloud switch Group, Huawei Technologies Co., Ltd. His research interests include vehicle safety, data-mining, cloud computing, and eco-driving.



YUAN SHENG received the bachelor's degree in engineering from Nagoya University, Japan, in 2016, where he is currently pursuing the master's degree in informatics. His research focuses on deep learning and driving behaviors analysis.

WENHU QIN received the Ph.D. degree from Southeast University, Nanjing, China, in 2005. Since 1997, he has been on the faculty of with the School of Instrument Science and Engineering, Southeast University, where he is currently a Professor. He directs the vehicle safety and virtual reality laboratory at Southeast University. He has authored or co-authored over 30 journal papers, 10 conference papers, and a book. He holds three patents. His research interests include vehicle safety, virtual reality, crowd simulation, and road traffic accident reconstruction.

CHIYOMI MIYAJIMA received the B.E., M.E., and Ph.D. degrees in computer science from the Nagoya Institute of Technology, Japan, in 1996, 1998, and 2001, respectively. From 2001 to 2003, she was a Research Associate with the Department of Computer Science, Nagoya Institute of Technology. She was an Assistant Professor with the Graduate School of Information Science, Nagoya University, Japan, from 2003 to 2016. She was an Associate Professor with the Institutes of Innovation for Future Society, Nagoya University, from 2016 to 2018. Since 2018, she has been an Associate Professor with Daido University, Nagoya, Japan. Her research interests include the analysis and the modeling of driver behavior.

KAZUYA TAKEDA (SM'09) received the B.E.E., M.E.E., and Ph.D. degrees from Nagoya University, Japan, in 1983, 1985, and 1994, respectively. Since 1985, he has been with the Advanced Telecommunication Research Laboratories and KDD R&D Laboratories, Japan. In 1995, he started a research group for signal processing applications at Nagoya University. He is currently a Professor with the Institutes of Innovation for Future Society, Nagoya University. His main focus is investigating driving behavior using data centric approaches, utilizing signal corpora of real-driving behavior. He is also a member of the Board of Governors of the IEEE Intelligent Transportation Systems Society.