

## Research Article

# Cognitive Distraction State Recognition of Drivers at a Nonsignalized Intersection in a Mixed Traffic Environment

Qiang Hua <sup>1</sup>, Lisheng Jin <sup>1,2</sup>, Yuying Jiang <sup>3</sup>, Ming Gao <sup>4</sup> and Baicang Guo<sup>1</sup>

<sup>1</sup>Transportation College, Jilin University, Changchun 130022, China

<sup>2</sup>School of Vehicle and Energy, Yanshan University, Qinhuangdao 066004, China

<sup>3</sup>Department of Ophthalmology, China-Japan Union Hospital, Jilin University, Changchun 130022, China

<sup>4</sup>Laboratory of Automotive Safety and Energy, Department of Automotive Engineering, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Yuying Jiang; [jiangyy@jlu.edu.cn](mailto:jiangyy@jlu.edu.cn)

Received 17 October 2020; Revised 21 January 2021; Accepted 23 February 2021; Published 4 March 2021

Academic Editor: Songtao Lv

Copyright © 2021 Qiang Hua et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Distracted driving has become a growing traffic safety concern. With advances in autonomous driving and connected vehicle technology, a mixture of various types of intelligent vehicles will become normal in the near future, while more factors that may cause driver cognitive distraction are emerging. However, there are rarely studies on distracted driving in mixed traffic environments. To fill this gap, we conducted a natural driving experiment with three representative events at a nonsignalized intersection in a mixed traffic environment and proposed a novel method of identifying cognitive distraction based on bidirectional long short-term memory (Bi-LSTM) with attention mechanism. Forty participants were recruited for each event, who completed three different cognitive distraction experiments induced by three different secondary tasks in contrast with a normal driving process when passing a nonsignalized intersection. Related driving performance and eye movement data were collected to train and test the Bi-LSTM with attention mechanism model. Compared with the support vector machine (SVM) model, its recognition accuracy rate is 94.33%, which is 3.83% higher than that of the SVM in the total event, which has reasonable applicability for distraction recognition in a mixed traffic environment. Potential applications of this model include distraction alarm and autonomous driving assistance systems, which could avoid road traffic accidents.

## 1. Introduction

Distracted driving has become a dominant cause of traffic accidents [1]. In the United States, 3450 people died of distracted driving in 2016 [2]. Another survey based on 1,367 drivers found that serious accidents caused by distracted driving account for 14% to 33% of traffic accidents [3]. Generally, distracted driving refers to the allocation of attention resources from driving to secondary tasks, causing the driver's remaining attention resources to be unable to guarantee the minimum attention demand required by the current driving environment, leading to a decline in driving performance capabilities [4]. When identifying whether a driver is in a distracted state, not only the driver's current behavioral characteristic status but also the complexity of the

traffic environment must be considered. With the application of intelligent network technology and advanced control algorithms in the automobile field [5], all types of intelligent vehicles are developing rapidly, and mixed traffic is bound to become normal, where manual driving vehicles (MVs), connected manual vehicles (CMVs) [6], automated vehicles (AVs) [7], and connected automated vehicles (CAVs) will simultaneously share the road. The complex traffic environment creates many incentives for distracted driving. Therefore, it is extremely crucial to study distracted driving considering mixed traffic environments [8]. The National Highway Traffic Safety Administration roughly divides distraction into visual distraction [9], operational distraction [10], and cognitive distraction [11]. Visual distraction, such as distracted visual attention during driving while focusing

on things not related to the driving area, can weaken driving safety, and visual distraction changes as the amount of time that the gaze point leaves the road ahead increases. Operational distraction refers to actions during driving that are not related to driving and have a detrimental effect on normal driving, such as smoking and manipulating the display while driving. Operational distraction changes as operational complexity increases.

Cognitive distraction means that the driver shifts his cognitive attention to secondary tasks that are not related to the driving task. Unlike visual distraction and operational distraction, when cognitive distraction occurs, the driver's eyes remain on the road ahead, and thus, this type of distraction is more concealed, such as daydreaming and hands-free cellphone conversation. Cognitive distraction changes with an increase in the cognitive load required for different secondary driving tasks. The driving information including the external driving information and in-vehicle driving information in a mixed traffic environment is more abundant than that in a traditional traffic environment, as a mixed traffic environment is particularly complicated [12], which is more likely to lead to cognitive distraction than a traditional traffic environment. The negative effects of cognitive distraction on driving in a mixed traffic environment are also more serious than those in a traditional traffic environment. Thus, detecting cognitive distraction is of paramount importance.

Approximately 18.1% of accidents each year are caused by distractions at traditional traffic environment intersections [13, 14]; however, traffic conditions are more complicated at a mixed traffic environment intersection. More attention is required from the driver in this environment than in other traffic environments [15], with a short stay and limited access to outside driving information. If cognitive distraction occurs, it will have a greater adverse impact on driving safety, increasing the probability of accidents. However, we found few studies on cognitive distraction in mixed traffic environments, especially at a nonsignalized intersection in a mixed traffic environment. The researchers [16] adopted a support vector machine (SVM) as a data mining method using driver eye movement data and driving performance data to develop a method for detecting cognitive distraction in real time [17]. The results showed that the SVM model can detect driver distraction with an average accuracy of 81.1%, which is better than that of a traditional logistic regression model [18]. A previous study [19] collected driving performance and driver eye movement data and applied a genetic selection algorithm with SVM-recursive feature elimination (SVM-RFE) to rank alternative indicators of driver distraction and obtain the order of importance for 29 candidate indicators. A driver distraction recognition model that is based on AdaBoost-genetic algorithm (GA)-backpropagation (BP) was established; using the driver distraction recognition model with the best feature index to identify a driver's state, the model's recognition accuracy rate was 95.09%, which indicated that the model could accurately identify the distraction state of a driver. A method for detecting driver cognitive distraction at stop-controlled intersections was proposed [20]. The method uses

the SVM-RFE algorithm to extract the optimal features from features constructed with driving performance and eye movement. After feature extraction, the SVM classifier is trained and cross-validated. The results showed that the SVM classifier based on the fusion of driving performance features and eye movement features achieved the best accuracy rate of 95.8% for stop-controlled intersections. The study [21] built an efficient model for detecting driver distraction and recognizing the type of distraction. Based on the active sensor Kinect and computer vision tools, the AdaBoost classifier and hidden Markov model were used to merge relevant information that is used for assessing driver inattention, and qualitative and quantitative results showed strong and accurate detection and recognition capacity. The researchers [22] proposed a real-time method for detecting a driver's cognitive and visual distraction using lateral driving performance measures, and radial basis probabilistic neural networks (RBPNNs) were adopted to construct classification models. As a result, the RBPNN model using the standard deviation of the lane position and steering wheel reversal rate could be an effective distraction detector with easy-to-obtain and inexpensive inputs. All of the above studies were based on traditional traffic scenarios and did not involve mixed traffic scenarios.

To fill this gap, this study proposes a method based on bidirectional long short-term memory (Bi-LSTM) with attention mechanism [23] to identify driver cognitive distraction at a nonsignalized intersection in a mixed traffic environment [24]. However, limited by actual conditions, complex mixed traffic scenarios cannot be completely reproduced. This paper studies cognitive distraction in three typical situations that occur at a nonsignalized intersection in a mixed traffic environment. These three situations are defined as three different events. In each event, the ego vehicle is the only vehicle that drives in the longitudinal lane, which is a CMV driven by the participant. The types of vehicles that approach from the lateral lane are different in each event; if a connected vehicle is approaching from the lateral lane, the ego vehicle will communicate with it to obtain the driving information of the lateral vehicle and eventually meet it at the intersection. In the driving process of each event, three different secondary tasks are arranged independently in turn, and if the participant of the ego vehicle is assigned a secondary task, this is considered to be distracted driving, while no assigned task is considered to be normal driving. To create the real effect of cognitive distraction in driving experiments, this study arranges a method in which a specific secondary task must be completed within a specified time to cause distracted driving. There are three types of secondary tasks (one-back task, two-back task, and clock task), each of which is easy to implement and requires different cognitive strengths. These tasks can be randomly arranged and repeated multiple times. The resulting distraction is appropriate.

Multiple real vehicle experiments are performed to collect driving performance data for distracted driving at nonsignalized intersections in mixed traffic environments, such as vehicle longitudinal and lateral control data as well as eye movement data from an eye tracker; the driver's

physiological state information is collected by a BIOPAC physiological recorder. Then, all of the data are subjected to fusion and preprocessing [25], and the optimal feature selection method based on the gray rough set theory is used to extract the optimal feature subset. Ten evaluation feature indices were extracted from 27 candidate feature indices and ranked based on the fuzzy analytic hierarchy process (FAHP). Different combinations of evaluation features were input into the Bi-LSTM with attention mechanism model to obtain the optimal number of feature combinations. Compared to the SVM model, the results show that the Bi-LSTM with attention mechanism model can reliably recognize the distracted driving behaviors of the participants with the best accuracy rate of 94.33% in the total event, which is 3.83% higher than that of SVM in the total event, and thus, it is suitable for distraction recognition in a mixed traffic environment. The findings of this study can be applied to distracted driving alarms for autonomous driving assistance systems.

## 2. Experimental System and Methods

**2.1. Participants.** This study was approved by the Ethics Committee of the China-Japan Union Hospital of Jilin University. A total of 40 drivers who were in satisfactory physical condition, had normal vision, and held legal C1 driver's licenses were recruited for the study. The recruited drivers included 20 males and 20 females, and their occupations included students and faculty members of Jilin University and urban office workers. Informed consent was obtained from the participants before the data collection stage of the study. The drivers were 21–57 years old, with a mean age of 29 years old and a standard deviation of 7.40 years. The driving experience of the participants ranged from 1 to 20 years, and the average length of driving experience was 5 years and the standard deviation was 3.46. The average total driving mileage was 39.2 thousand kilometers and the standard deviation was 14.6. None of the participants exhibited any adverse effects, such as drug effects, fatigued driving, or drunk driving, during the test. Table 1 shows the group personal statistics.

**2.2. Experimental Equipment.** The equipment utilized in the driving experiment includes connected vehicles, VAG-COM Diagnostic System (VCDS) equipment, a BIOPAC physiological recorder, and the Tobii Pro Glasses 2 eye tracker, as shown in Figure 1. The connected vehicles could communicate with each other in real time, which allows the driving information of other connected vehicles, such as the current speed of a vehicle, the distance from the vehicle, and the distance from the vehicle to the intersection, to be obtained. The participant's driving performance data are directly collected by VCDS equipment, and the driving performance data are output in the CSV format.

The participant's physiological state information is collected by a BIOPAC physiological recorder. In this experiment, electroencephalography (EEG) 100C amplifiers, 2 LEAD110 shielded wires, 2 LEAD100 unshielded wires, and

4 disposable patch electrodes were used to collect EEG signals, and the corresponding AcqKnowledge 4.2 physiological signal processing software was utilized to preprocess the original physiological signal data and extract the valid data.

The Tobii Pro Glasses 2 eye tracker has a sampling rate of 50 Hz and four eye-tracking cameras, which can obtain a driver's blink frequency, fixation direction, saccade frequency, degree of eyelid opening, and other information [26]. Additionally, the tracker can obtain the information of the road ahead and visually present the driver's fixation point and fixation track in a scene by superimposition with the video of the driver information acquisition system. Eye-tracking data acquired by Tobii Pro Glasses 2 can be synchronized with a wide range of physiological data, including EEG [27], EMG, motion detection, breathing, and heart rate.

**2.3. Experimental Design.** A nonsignalized mixed traffic environment intersection is designed in this paper. The location is on the Nanling Campus of Jilin University, and the road is a two-way two-lane road. Participants were asked to drive CMVs in the longitudinal lane, which is the only vehicle that drives in the longitudinal lane. In the lateral lane, there are MVs, CMVs, AVs, and CAVs, and the maximum speed does not exceed 50 km/h. The appearance of vehicles in lateral lanes was triggered when the ego vehicle passed 150 m before the center of the intersection, and connected vehicles could communicate with the ego vehicle. Information was reported to the participant in the ego vehicle in the form of voice reminders, and the content was the speed and location of the vehicle in the lateral lane. Since the types of vehicles that approach from the lateral lane are different, this paper divides the scenarios into three different types of events, as shown below:

- (i) Event 1: the first vehicle approaching from the right side of the lateral lane is an MV, and the second vehicle is an AV; there is no communication between the lateral vehicle and the ego vehicle in the longitudinal lane
- (ii) Event 2: the first vehicle approaching from the right side of the lateral lane is a CMV, and the second vehicle is an MV; there is communication between the CMV in the lateral lane and the ego vehicle in the longitudinal lane
- (iii) Event 3: the first vehicle approaching from the right side of the lateral lane is a CMV, and the second vehicle is a CAV; there is communication between the two vehicles in the lateral lane and the ego vehicle in the longitudinal lane

The time headway between two vehicles in the lateral lane is randomized from 1 s to 3 s. All vehicles in the lateral lane are driven by professionals. When the ego vehicle is 150 m before the center of the intersection, the appearance of vehicles in the lateral lane and the secondary tasks are triggered. Eventually, the ego vehicle in the longitudinal lane meets the vehicles in the lateral lane at the intersection and then drives away from the intersection.

TABLE 1: Main effect test table.

Gender	Sample size	Mean age (years)	Average mileage (thousand km)	SD of mileage (thousand km)
Male	20	30	45.7	17.8
Female	20	28	32.7	13.4



FIGURE 1: Vehicles and information collection equipment.

The process of intersection crossing is extracted from 150 m before the ego vehicle crosses the stop line of the intersection to 20 m after the center of the intersection in each event [28], and each event contains normal driving and driving with three different secondary tasks, respectively.

Driving with the secondary tasks is considered distracted driving, and driving without the secondary tasks is considered normal driving. All the durations of the secondary tasks exceeded the time of the extracted event data, which ensures that distracted driving covered the entire target event. The above measures were repeated to extract data under each representative important event, and the number of normal driving instances collected for each event was the same as the number of every secondary task driving instance collected in each event.

A 15-minute test drive and secondary task training were conducted with participants before the formal experiment to ensure that the participants were familiar with the entire experimental process. The experimental scenario is shown in Figure 2. There are three nonsignalized intersections that cross the longitudinal lane, each participant was required to drive the experimental road six times to encounter 18 intersections under each driving situation, and four driving situations (one-back, two-back, clock tasks, and normal driving) were involved in each event for the 40 participants. Therefore, a total of 2880 sets of data were collected from all participants in each event, and 1800 sets of data were selected from the 2880 sets of data, including 720 sets for normal driving and 1080 sets for distracted driving. The total event is the sum of the three events and includes 720 sets for normal driving and 1080 sets for distracted driving (240 sets of normal driving data and 360 sets of distracted driving data were selected from each event).

**2.4. Cognitive Distraction Generation.** When the ego vehicle was 150 m before the center of the intersection, a test

assistant in the ego vehicle arranged secondary tasks for the participants, and these tasks were divided into three types of secondary tasks [29]: one-back, two-back [30], and clock tasks. These chosen surrogate secondary tasks consumed the drivers' cognition and easily distracted them, and they were also easy to implement. The correspondingly scheduled secondary tasks required the participants to answer questions. All of the durations of the question-and-answer process exceeded the time of the extracted event data; there was an interval of 2 s between two questions. If the participant could not answer, they proceeded to the next question; the details of the secondary tasks are as follows:

- (i) One-back task: the assistant says a number, and participants say the preceding number of the number given by the assistant. There is no need to answer the first question (the value ranges from 0 to 9).
- (ii) Two-back task: the assistant says a number, and participants say the preceding two numbers of the number given by the assistant. There is no need to answer the first two questions (the value ranges from 0 to 9).
- (iii) Clock task: the assistant says a random time for participants to determine whether the angle between the hour and minute hands is an acute angle (1:00–12:59).

If a problem was encountered during an experiment, the participants restarted the actual driving to ensure the validity of the experimental data. During the experiment, except for the secondary task set in this experiment, no other human factors interfered with participants to ensure the validity of the data, and the time of each experiment was generally consistent, between 8–11 am and 2–5 pm. The secondary tasks are shown in Figure 3.

**2.5. Cognitive Load Assessment.** The value of the EEG index  $R(\theta + \beta)$  was applied to evaluate the cognitive load [31]. During an experiment, the participant wore a BIOPAC physiological recorder, which consisted of many modules. The EEG module can collect the driver's EEG signals, export collected data information in CSV format, import the original EEG data into MATLAB for noise reduction, and then extract the EEG indicator  $R(\theta + \beta)$  value [32]. SPSS is used to statistically analyze the EEG index  $R(\theta + \beta)$  values in different secondary tasks. The purpose of analyzing the EEG index  $R(\theta + \beta)$  values in different secondary tasks is to verify that each secondary task increased the cognitive load of the participant [33] and then caused a driving distraction, which verifies the feasibility of the designed secondary tasks, and that the cognitive load level required for each secondary task

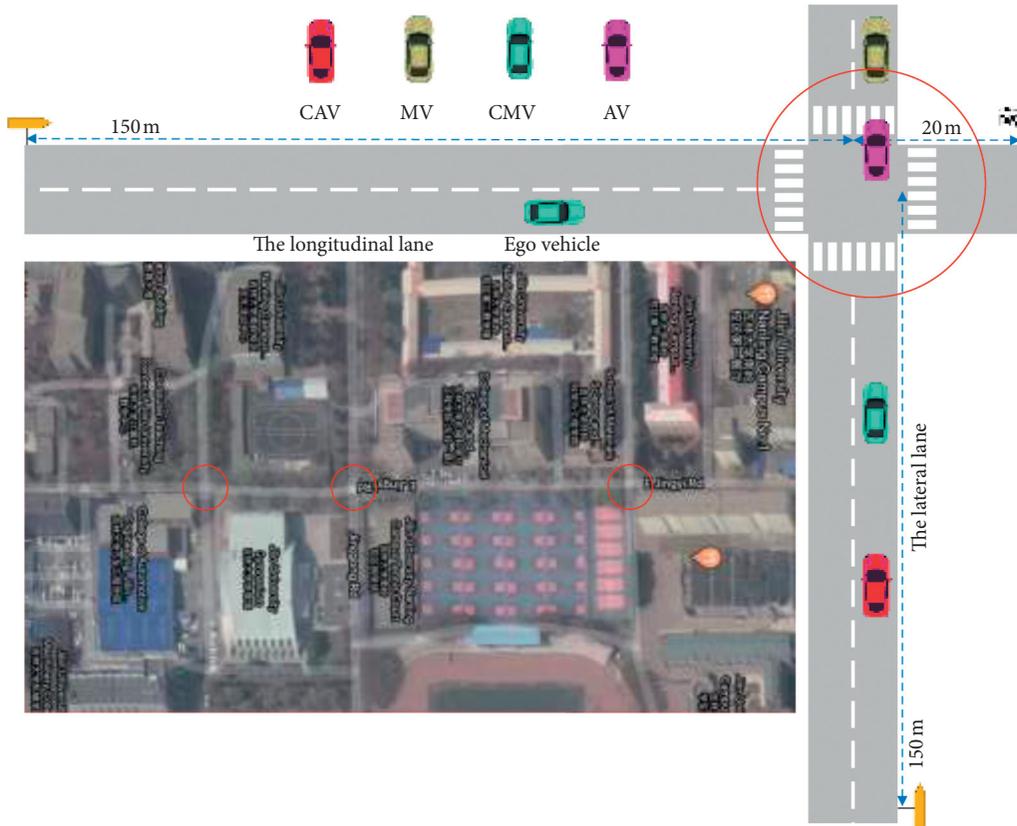


FIGURE 2: Illustration of experimental scenario.

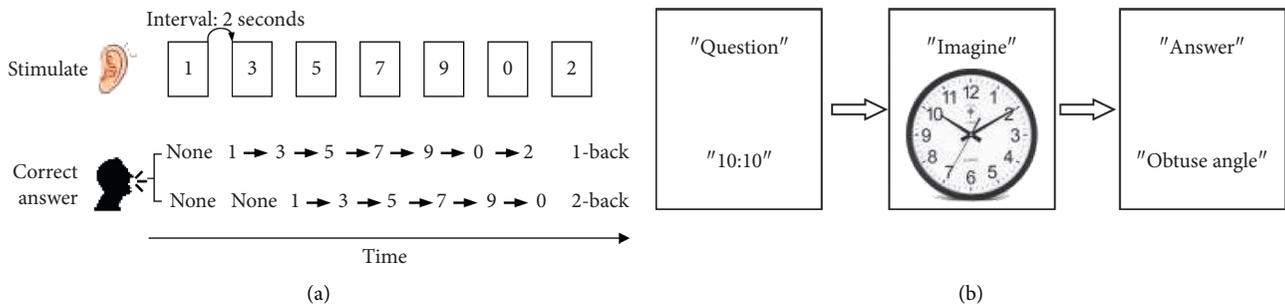


FIGURE 3: Illustration of the cognitive secondary tasks. (a) N-back task. (b) Clock task.

was different. This approach also verified the feasibility of research on distracted driving under different levels of secondary tasks. The EEG index extraction process is shown in Figure 4.

**2.6. Data Description.** Data collection was performed according to the experimental design process, that is, data were collected for the entire process of participants driving the vehicle 150 m before the intersection stop line to 20 m after the center of the intersection. Basic characteristics that were collected include the driving performance features and eye movement features, which provided a basis for calculating the average, variance, and correlation analysis of

various parameters. To extract the most important features for cognitive distraction detection, the candidate feature set was divided into two categories.

**2.6.1. Driving Performance-Related Features.** Driving performance-related features were calculated based on driving performance features, including lateral control performance features and longitudinal control performance features. The corresponding driving performance-related features are shown in Table 2.

**2.6.2. Eye Movement Features.** The eye movement data were collected in real time through the Tobii Pro Glasses 2 eye

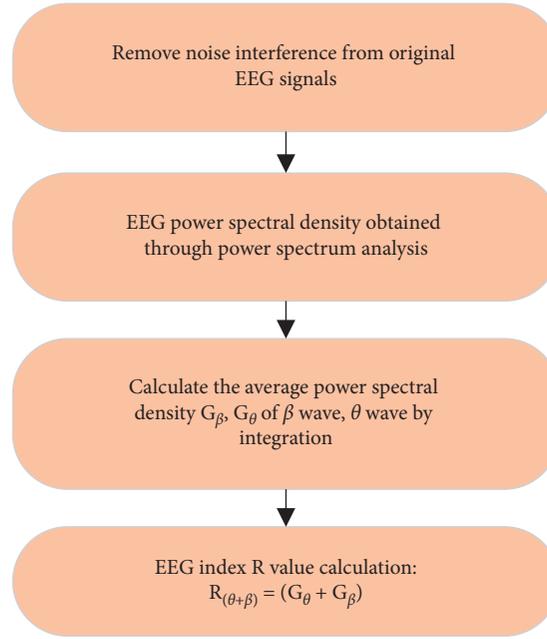


FIGURE 4: EEG index extraction process.

TABLE 2: Driving performance-related features.

Dependent variable	Variable description	Unit
Min_Dec	Minimum deceleration	m/s <sup>2</sup>
SWA_SD	Standard deviation of steering wheel angle	°
Accelerator pedal $H_b$	Accelerator pedal opening	cm
SWA_Mean	Mean steering wheel angle	°
Distance $H_p$	Distance from vehicle to intersection	m
SP_SD	Standard deviation of speed	m/s
SP_Mean	Mean speed	m/s
LA_Mean	Mean lateral acceleration	m/s <sup>2</sup>
Braking position $H_p$	Brake pedal opening	cm
SE_Mean	Mean steering entropy, which measures the smoothness of steering control without turning behavior to quantify the workload imposed on drivers	—

tracker worn by participants. The eye data collection time was synchronized with the driving performance data collection time. The corresponding eye movement features are shown in Table 3.

## 2.7. Feature Screening and Model Construction

**2.7.1. Screening the Optimal Features Based on Gray Rough Set Theory.** Rough set theory and gray system theory are both effective mathematical theories for dealing with uncertainty problems. Gray clustering is used to replace the equivalent relationship in rough set theory to form gray rough set theory. This theory can deal with uncertain information from different aspects. To obtain a scientific and

reasonable evaluation index system of optimal features, gray rough set theory is used to screen feature evaluation indices [34]. Using gray cluster analysis theory and rough set theory, 10 evaluation feature indices were screened from 27 candidate feature indices [35, 36]. The corresponding candidate features are shown in Tables 2 and 3.

**2.7.2. Weight Calculation of the Optimal Features Using the FAHP.** Since 10 evaluation feature indices were selected, a streamlined optimal feature evaluation index system was built. On this basis, the FAHP [37] was used to determine the weight values of the optimal feature evaluation index, and the extracted optimal feature evaluation index was sorted. It is provided in Tables 4 and 5. This approach lays a

TABLE 3: Eye movement features.

Feature	Variable description	Unit
Blink F	Blink frequency	Hz
Fixation P_SD	Standard deviation of fixation point position in the horizontal direction	°
Saccade F	Saccade frequency	Hz
SacAmp_Mean	Mean saccade amplitude	°
Blink T_Mean	Mean blink time	ms
Blink T_SD	Standard deviation of blink time	ms
Fixation T_Mean	Mean fixation time	ms
Fixation T_SD	Standard deviation of fixation time	ms
Pupil_SD	Standard deviation of pupil diameter	mm
Pupil_Mean	Mean pupil diameter	mm
SacAmp_SD	Standard deviation of saccade amplitude	°
SacPSP_Mean	Mean saccade peak speed	°/s
SacPSP_SD	Standard deviation of saccade peak speed	°/s
SacSP_Mean	Mean saccade speed	°/s
SacSP_SD	Standard deviation of saccade speed	°/s
SaccadeT_Mean	Mean of saccade time	ms
SaccadeT_SD	Standard deviation of saccade time	ms

TABLE 4: Feature-level division.

First-level feature indices	Weights	Secondary-level feature indices	Weights
Eye movement features	0.631	Blink F	0.301
		Fixation P_SD	0.286
		Saccade F	0.175
		SacAmp_Mean	0.162
		Pupil_Mean	0.157
		Pupil_SD	0.143
Driving performance-related features	0.502	Blink T_SD	0.126
		SP_Mean	0.347
		SWA_SD	0.262
		SE_Mean	0.165

TABLE 5: Optimal feature index ranking.

Rank	Optimal feature index	Variable description
1	SP_Mean	Mean speed
2	Blink F	Blink frequency
3	Fixation P_SD	Standard deviation of fixation point position in the horizontal direction
4	SWA_SD	Standard deviation of steering wheel angle
5	Saccade F	Saccade frequency
6	SE_Mean	Mean steering entropy
7	SacAmp_Mean	Mean saccade amplitude
8	Pupil_Mean	Mean pupil diameter
9	Pupil_SD	Standard deviation of pupil diameter
10	Blink T_SD	Standard deviation of blink time

foundation for the next step in understanding the influence of the different numbers of optimal features on the recognition model.

**2.7.3. LSTM Network.** LSTM is a recurrent neural network (RNN), which achieves the same satisfactory performance as the standard RNN when processing times series data. However, compared to the standard RNN, LSTM can remember important features or moments in long-term

historical time series data, which introduce forget gate to prevent gradient explosion and gradient vanish. The standard LSTM consists of three parts: input gate  $\mathbf{i}_t$ , output gate  $\mathbf{o}_t$ , and forget gate  $\mathbf{f}_t$ . The calculation results can be output via a series of calculations in the hidden state inside the cell, where  $\sigma$  represents a logistic sigmoid function,  $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_o, \mathbf{W}_c$  represents the cyclic weight, and  $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c$  represents the corresponding bias. The internal calculation process of the LSTM cell is as follows.

The first step is to select the information that will be discarded in the nerve cell through a sigmoid layer of the forgetting gate. The gate will draw the input  $\mathbf{x}_t$  of the current layer and the output  $\mathbf{h}_{t-1}$  of the hidden layer from its upper layer and then output a value between 0 and 1 to the cell state  $\mathbf{C}_{t-1}$ . The formula is provided as follows:

$$\mathbf{f}_t = \sigma(\mathbf{w}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f). \quad (1)$$

The second step is to determine what new information is stored in the cell state, which is composed of two parts. First, a sigmoid layer of the input gate determines what information should be updated. A  $\tanh$  layer creates the new candidate state  $\tilde{\mathbf{C}}_t$ , which is employed to update the next state, as shown in the following formula:

$$\begin{aligned} \mathbf{i}_t &= \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \\ \tilde{\mathbf{C}}_t &= \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c). \end{aligned} \quad (2)$$

The third step is to update the old cell state  $\mathbf{C}_{t-1}$  to the new cell state  $\mathbf{C}_t$ . First,  $\mathbf{C}_{t-1}$  and  $\mathbf{f}_t$  are multiplied to discard unimportant information, and second,  $\mathbf{i}_t * \tilde{\mathbf{C}}_t$  is added to generate the new state  $\mathbf{C}_t$  of the nerve cell, as illustrated in the following formula:

$$\mathbf{C}_t = \mathbf{f}_t * \mathbf{C}_{t-1} + \mathbf{i}_t * \tilde{\mathbf{C}}_t. \quad (3)$$

Last, the output value is determined based on the cell state  $\mathbf{C}_t$ . First, a sigmoid layer of the output gate is used to determine what important information should be output. Second,  $\tanh$  is employed to process the cell state  $\mathbf{C}_t$ , and the result is multiplied by the output of  $\mathbf{o}_t$  to obtain the last output  $\mathbf{h}_t$ . This process is shown as follows:

$$\begin{aligned} \mathbf{o}_t &= \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o), \\ \mathbf{h}_t &= \mathbf{o}_t * \tanh(\mathbf{C}_t). \end{aligned} \quad (4)$$

**2.7.4. Bi-LSTM Model with Attention Mechanism.** Bi-LSTM is composed of forward LSTM and backward LSTM. The hidden layer state  $\mathbf{h}_t$  at time  $t$  is derived from the weighted sum of the forward hidden layer states  $\mathbf{h}_t$ ; the backward hidden layer states  $\mathbf{h}_t$ ,  $\mathbf{U}_t$ , and  $\mathbf{V}_t$  are the corresponding weights; and  $\mathbf{b}_t$  represents the bias of the hidden layer state at time  $t$ . The relationship among the distraction features can be learned more comprehensively to further improve the recognition accuracy. Applying an attention mechanism in the Bi-LSTM framework enables more attention to be paid to the specific input feature sequence, which eliminates redundant features due to the attention weight of features and captures the key feature information. Thus, the validity and accuracy of the recognition model are improved. The Bi-LSTM framework-introduced attention mechanism is illustrated in Figure 5.

The details of the attention mechanism are provided as follows:

$$\vec{\mathbf{h}}_t = \text{LSTM}(\mathbf{x}_t, \vec{\mathbf{h}}_{t-1}),$$

$$\overleftarrow{\mathbf{h}}_t = \text{LSTM}(\mathbf{x}_t, \overleftarrow{\mathbf{h}}_{t-1}),$$

$$\mathbf{h}_t = \mathbf{U}_t \vec{\mathbf{h}}_t + \mathbf{V}_t \overleftarrow{\mathbf{h}}_t,$$

$$\mathbf{e}_t = \mathbf{F}(\mathbf{h}_t), \quad (5)$$

$$\mathbf{W}_t = \frac{\exp(\mathbf{e}_t)}{\sum_{i=1}^n \exp(\mathbf{e}_i)},$$

$$\mathbf{a} = \sum_{i=1}^n \exp(\mathbf{W}_t \mathbf{h}_t),$$

$$\mathbf{y}_j = \text{softmax}(\mathbf{a} \mathbf{W}_j + \mathbf{b}_j),$$

where  $\mathbf{F}$  is a learning function, which is used to calculate the weight  $\mathbf{W}_t$  of the output vector  $\mathbf{h}_t$  of the LSTM layer at time  $t$ , and the final feature representation vector  $\mathbf{a}$  is obtained by weighting. Finally, the recognition result of cognitive distraction is transmitted through the softmax layer.  $\mathbf{W}_j$  and  $\mathbf{b}_j$  are the weight coefficient matrix and the corresponding bias in the output layer, respectively.

**2.7.5. Model Construction and Identification Process.** Cognitive distraction state recognition can be regarded as a time series modeling and prediction problem. Feature indices are extracted through the continuous data flow of time series in cognitive distraction and normal driving. Model inputs are the eigenmatrix, which consists of evaluation feature index vectors, and the eigenvectors are composed of the time series of each evaluation feature index 0-T s.

The identification process of the distraction state recognition model of drivers based on Bi-LSTM with attention mechanism at a nonsignalized intersection in a mixed traffic scenario is described as follows:

- (i) Use the “leave one out method” [38] to divide the test set and training set of the model. During each training, select the current single sample as the test set and the remaining samples as the training set to ensure that the training set includes normal driving and distracted driving states in each event.
- (ii) To avoid the influence of different unit cell dimensions between different parameters and eliminate the differences between the indices, the min-max normalization method is used to normalize the sample data:

$$\mathbf{x}' = \frac{(\mathbf{x} - \mathbf{x}_{\min})}{(\mathbf{x}_{\max} - \mathbf{x}_{\min})}, \quad (6)$$

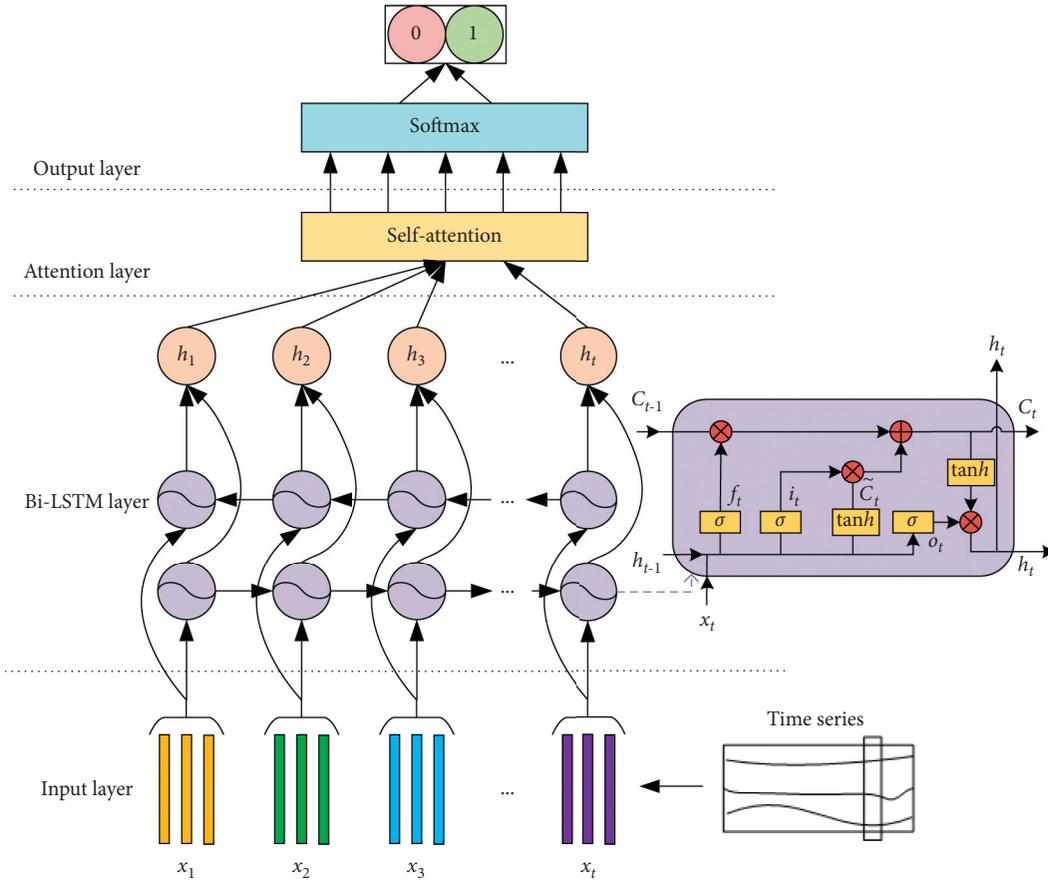


FIGURE 5: Bi-LSTM framework with the attention mechanism.

where  $\mathbf{x}'$  is the normalized value,  $\mathbf{x}$  is the original value of the sample,  $\mathbf{x}_{\min}$  is the minimum value of the sample, and  $\mathbf{x}_{\max}$  is the maximum value of the sample.

- (iii) The first  $n$  (4–10) feature index vectors in Table 5 are selected as the input eigenmatrix for driver distraction recognition in each event, which are input into the model to obtain the performance of driver distraction recognition models with a different number of feature indices.
- (iv) The expert experience method is applied to set the parameter of the Bi-LSTM, employing the default parameter setting to run the algorithm, where the number of hidden units is 100, the initial learning rate is 0.1, the dropout rate is 0.5, and the max number of epochs is 60. After repeated experiments on the validation set, the loss tends to be stable, and the optimal adjustment parameters are obtained. An Adam optimizer is adopted in the network, with a learning rate of 0.01 and an attenuation of 0.9. The number of hidden units per layer is 128; the max number of epochs is 90, and the dropout rate is 0.4.
- (v) Model performance evaluation: for the test set data, accuracy, precision, recall, and  $F_1$  values were

utilized to evaluate the model performance, which are introduced as follows:

$$\begin{aligned}
 P &= \frac{TP}{TP + FP}, \\
 R &= \frac{TP}{TP + FN}, \\
 A &= \frac{TP + TN}{TP + FN + FP + TN}, \\
 F_1 &= \frac{2TP}{2TP + FN + FP},
 \end{aligned} \tag{7}$$

where TP is the number of instances labeled 1 that are predicted to be 1, FN is the number of instances labeled 1 that are predicted to be 0, FP is the number of instances labeled 0 that are predicted to be 1, and TN is the number of instances labeled 0 that are predicted to be 0. 1 = distracted driving, 0 = normal driving.  $P$  indicates the proportion of positive examples judged to be true.  $R$  is the ratio of positive examples identified as true to the total positive examples in the classifier.  $A$  is the ratio of the number of instances correctly classified to the total

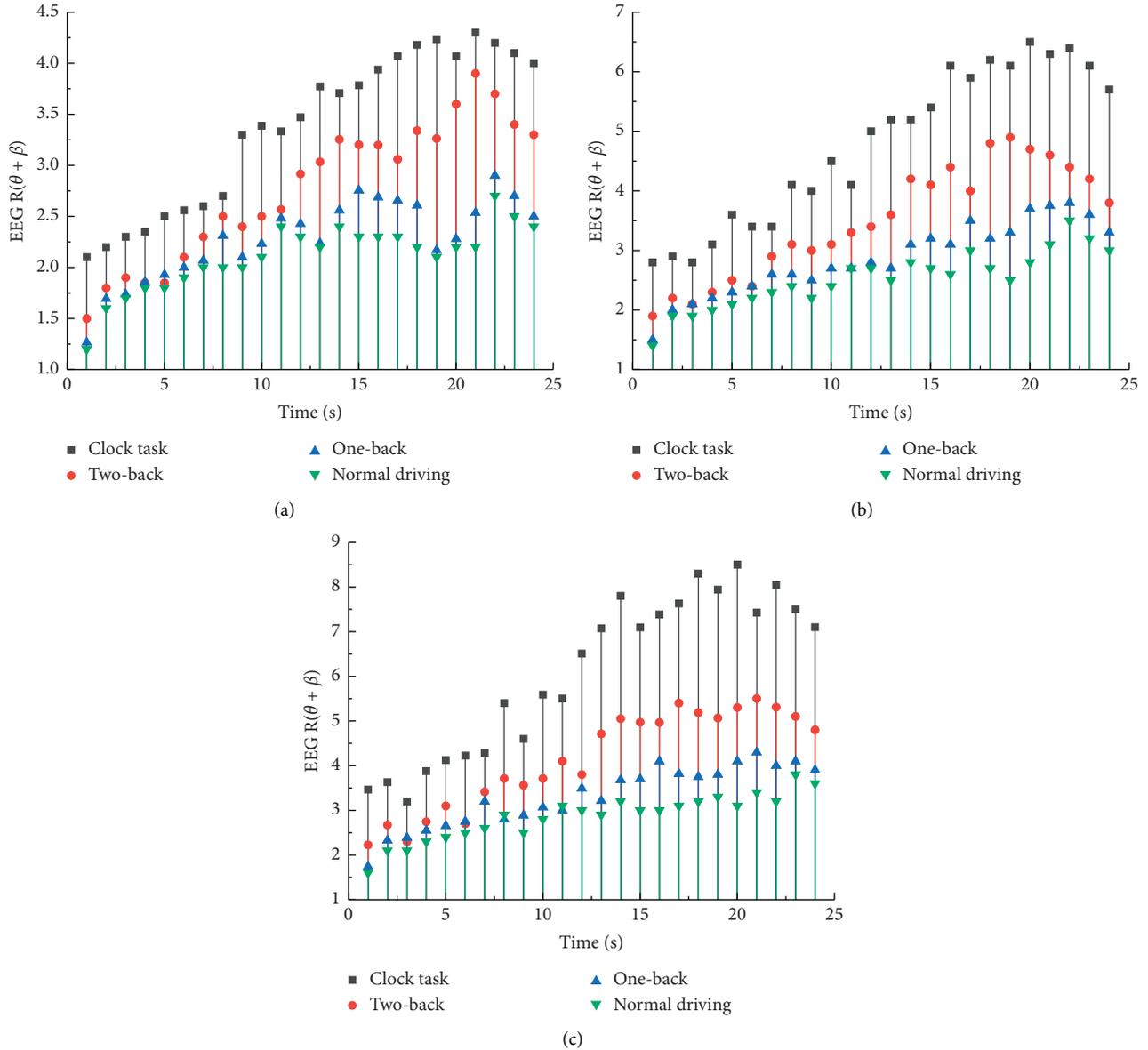


FIGURE 6: EEG index  $R(\theta + \beta)$  values of the three events in different secondary tasks, (a) participant's cognitive load graph for event one, (b) participant's cognitive load graph for event two, and (c) participant's cognitive load graph for event three.

number of instances.  $F_1$  is the harmonic average of  $P$  and  $R$ .

**2.7.6. Description of the SVM Method.** The SVM was first proposed by Cortes and Vapnik in 1995 [39], which has many unique advantages in solving small sample, nonlinear, and high-dimensional pattern recognition. SVMs are based on a statistical learning technique and can be used for pattern classification and inference of nonlinear relationships between variables. This method has been successfully applied in the field of recognition and detection [40].

The labeled binary class data  $\mathbf{B}_t = \{(\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_i^T, \mathbf{y}_i)\}$  is the input eigenmatrix, where  $\mathbf{x}_i$  is the eigenvector, which is composed of the first  $n$  (4–10) features in Table 5,  $T$  is the time series length of 0-Ts, and  $\mathbf{y}_i$  is a class indicator with a

value of either 0 or 1. In this paper, the radial basis function (RBF) kernel function is selected via many tests, and the optimal parameter was determined, where  $C = 16.57$  and  $\gamma = 53.61$ .

### 3. Results

The EEG signals of the driver were collected by a BIOPAC physiological recorder, and the EEG index  $R(\theta + \beta)$  values that represented the cognitive load were extracted. SPSS software was employed to statistically analyze the EEG index  $R(\theta + \beta)$  values of the collected data. Cognitive load graphs of one participant in the three events are shown in Figure 6.

It can be seen that when the secondary tasks were added, the cognitive load required by the participants increased accordingly. The higher the EEG index  $R(\theta + \beta)$  value is, the



FIGURE 7: Continued.



FIGURE 7: Distribution of the hot spots in the area of interest of the participant: (a) normal driving in event one, (b) distracted driving (one-back) in event one, (c) distracted driving (two-back) in event one, (d) distracted driving (clock task) in event one, (e) normal driving in event two, (f) distracted driving (one-back) in event two, (g) distracted driving (two-back) in event two, (h) distracted driving (clock task) in event two, (i) normal driving in event three, (j) distracted driving (one-back) in event three, (k) distracted driving (two-back) in event three, and (l) distracted driving (clock task) in event three.

higher the corresponding participant's cognitive load is. Figure 6 also shows that each secondary task increased the cognitive load of the participant to be higher than the cognitive load required in normal driving conditions; this situation will cause distracted driving. This also verifies the rationality of considering secondary tasks [41]. The EEG index  $R(\theta + \beta)$  value of event three was higher than that of event one and event two for the same secondary task, which indicates that the cognitive load required for the participant to drive in event three was higher than the cognitive load required for the participant to drive in event one and event two because the scenario of event three is more complex than that of the remaining two events, which produces a higher cognitive load. When the clock task was conducted in each event, the participant experienced the greatest cognitive load. The two-back task was second, and the one-back task corresponded to the smallest cognitive load, indicating that the clock task was the most difficult, requiring a larger cognitive load. In addition, the cognitive load in all tasks has an increasing trend with the passage of time, and it begins to decrease after reaching its peak. This occurred because as the experimental vehicle approached the nonsignalized intersection, the driving environment became more complicated, and the required cognitive load significantly increased. When vehicles leave the intersection, the traffic environment becomes simple, and the cognitive load required is reduced.

Figure 7 shows the distribution of the hot spots in the areas of interest of the participants during normal driving and distracted driving. In the corresponding events, under normal driving conditions in each event, the participant's fixation points were concentrated on the vehicle approaching from the lateral lane of the intersection. When one-back and two-back tasks were employed, the participant's fixation points were scattered. However, most of the fixation points were concentrated on the front driving area, and the participant's fixation points were mostly concentrated directly above the front driving area when the clock task was employed. We argue that when the participant has a large cognitive load required to think about the question of the clock task, the participant looks above the front driving area and can no longer allocate his attention to the driving area of the intersection.

**3.1. Selected Feature Indices.** The weight values of the optimal features are shown in Table 4. The weight values of the eye movement features were greater than the weight values of the driving performance-related features in the first-level evaluation indices, indicating that the eye movement feature indices better reflected the distracted driving state than the driving performance-related feature indices.

The weight value of the SP\_Mean index was 0.347 in the second-level evaluation indices, which indicates that SP\_Mean was the most important evaluation index among all evaluation indices. The ranking of second-level evaluation indices was obtained according to the weight values. The ranking of optimal feature indices is shown in Table 5.

**3.2. Model Analysis Results.** Tables 6–9 present the results of distracted driver recognition based on the Bi-LSTM with attention mechanism. As shown in Table 6, as the number of input features increases, the distraction recognition accuracy rate continues to increase. When the first 8 feature indices were input in the total event, the recognition accuracy rate was the highest, with a value of 94.33%, and the precision value was 96.57%, the recall was 93.88%, and the  $F_1$  value was 95.21%. When the number of input feature indices exceeded eight, the recognition accuracy decreased and tended to stabilize. For distraction recognition in event one, the recognition accuracy rate continued to increase as the number of feature indices increased. When the first 8 feature indices were input, the recognition accuracy rate was the highest, with a value of 89.16%, and the precision value was 93.25%, the recall was 88.33%, and the  $F_1$  value was 91.71%. When more than 8 features were input, the recognition accuracy also tended to be stable.

It is not obvious from the results that the more the feature indices there are, the better the recognition effect. The optimal number of input feature indices for distraction recognition was 7 in event two and event three, and the recognition accuracies were 93% and 96.16%, respectively. The optimal number of input features in each event also varied, and event three was the most complicated scenario among the three events. Event two was second, but the optimal distraction recognition accuracy of event three was

TABLE 6: Model performance with different indicators in total event.

Number of features	Accuracy (%)	Precision (%)	Recall (%)	$F_1$ (%)	Event
4	87.66	92.30	86.66	89.39	Total event
5	89.50	93.54	88.61	91.01	
6	91.83	95.07	91.11	92.63	
7	92.33	95.37	91.66	93.48	
8	94.33	96.57	93.88	95.21	
9	93.50	95.92	93.05	94.49	
10	91.16	94.49	90.55	92.48	

TABLE 7: Model performance with different indicators in event one.

Number of features	Accuracy (%)	Precision (%)	Recall (%)	$F_1$ (%)	Event
4	83.66	89.45	82.50	85.83	Event one
5	85.33	90.70	84.16	87.31	
6	86.50	91.39	85.55	88.37	
7	88.83	92.96	88.05	90.44	
8	89.16	93.25	88.33	91.71	
9	88.83	92.86	88.05	90.44	
10	87.16	91.98	86.11	88.95	

TABLE 8: Model performance with different indicators in event two.

Number of features	Accuracy (%)	Precision (%)	Recall (%)	$F_1$ (%)	Event
4	87.43	92.11	86.45	89.39	Event two
5	89.33	93.27	88.61	90.88	
6	91.38	94.50	90.83	93.04	
7	93	95.68	92.50	94.06	
8	92.50	95.38	91.94	93.63	
9	91.16	94.49	90.55	92.48	
10	88.66	92.94	87.77	90.28	

TABLE 9: Model performance with different indicators in event three.

Number of features	Accuracy (%)	Precision (%)	Recall (%)	$F_1$ (%)	Event
4	89.66	93.56	88.85	91.16	Event three
5	91.16	94.49	90.55	92.48	
6	93.66	96.26	93.05	94.63	
7	96.16	97.26	95.83	96.77	
8	94.16	96.03	93.82	95.09	
9	92.50	95.38	91.94	93.63	
10	90.33	94.15	89.44	91.73	

the highest, indicating that the algorithm is applicable to scenarios with high complexity. Tables 6–9 also show that different numbers of input features produced different recognition accuracies for distracted driving in the same event. If too many optimal features are input, the accuracy of distraction recognition will tend to decline. Therefore, the number of input feature indices has a certain influence on distracted driver recognition.

**3.3. Comparison with the SVM in the Total Event.** The SVM algorithm was used for model accuracy comparison in the total event. Table 10 shows that the Bi-LSTM model with

attention mechanism outperformed the SVM. The recognition accuracy rate of the Bi-LSTM model is 94.33%, and the  $F_1$  value is 95.21%, which are higher than the recognition accuracy rates of SVM, whose recognition accuracy rate is 90.5% and  $F_1$  value is 91.91%. Figure 8 shows the receiver operating characteristic (ROC) curve of the two recognition models for the total event. The two curves display satisfactory results, which illustrates that while both models have acceptable recognition performance, the Bi-LSTM model is better than the SVM algorithm. The optimal number (8) of the input feature indices, which are the top eight features in the optimal feature index ranking, as shown in Table 5, is the same for the two models. From these results, which illustrate

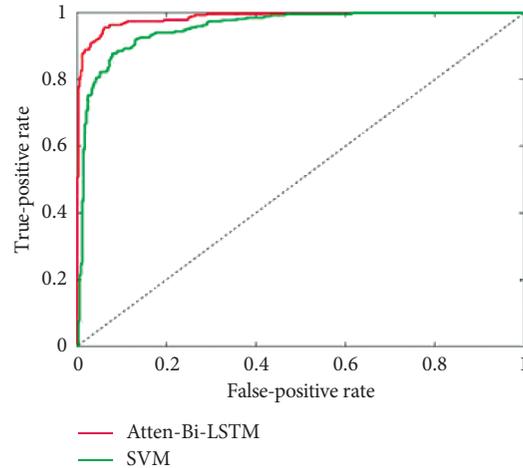


FIGURE 8: ROC curves of the identification by the two models in the total event.

TABLE 10: Comparison of distracted driving behavior recognition results of different algorithms.

Optimal number of features	Accuracy (%)	Precision (%)	Recall (%)	$F_1$ (%)	Type of model	Event
8	94.33	96.57	93.88	95.21	Atten-Bi-LSTM	Total event
8	90.5	93.91	90	91.91	SVM	

that the Bi-LSTM model was more accurate than the SVM model for distracted driving, we can draw a conclusion that the Bi-LSTM model has a reasonable effect on distraction recognition in mixed traffic environments.

#### 4. Discussion

Mixed traffic environments will be more complicated as more factors that induce driver cognitive distraction are introduced. To explore a driver cognitive distraction recognition method in a mixed traffic environment, this paper presents a method for identifying cognitive distraction based on Bi-LSTM with attention mechanism at a nonsignalized intersection in a mixed traffic environment, using the combinations of the optimal feature indices as inputs to the classifier.

The optimal feature screening method based on the gray rough set theory was used to extract the optimal feature indices from driving performance-related features and eye movement features. Twenty-seven candidate indices were reduced to 10, a scientific and reasonable evaluation index system of optimal features was established, and the FAHP was used to determine the weight values of the feature evaluation indices and rank the feature indices. The weight value of the eye movement features was the greatest in the first-level evaluation indices, and there was a large difference in the distribution of the hot spots in the area of interest of the participant during normal driving and distracted driving in event 1, event 2, and event 3. Figure 7 shows that the fixation points are relatively concentrated during normal driving, while the fixation points are relatively divergent during distracted driving. The results indicate that the eye

movement index plays a crucial role in distraction recognition in mixed traffic environments.

The cognitive load required by participants for driving with secondary tasks is greater than the cognitive load required for normal driving, and this finding verifies the rationality of considering secondary tasks. Observing the cognitive load of the participant with the same secondary task in these three events, Figure 6 shows that the cognitive load required by the participant in event 3 was the highest, which indicates that the more complex the scenario is, the greater the driving load required for the participant, and a complex scenario more easily causes distracted driving. When the participants were given different secondary tasks in each event, the cognitive load required for the clock task was higher than those for the two-back and one-back tasks, and this result indicates that the clock task is the most difficult and the most likely to cause cognitive distraction. As shown in Tables 6–9, the recognition accuracy rate varies with the number of optimal feature combinations. After the best number of optimal feature combinations is reached, the recognition accuracy rate gradually decreases and tends to be stable. The driving scenario of event 3 is the most complicated, and the driving scenario of event 2 is the second most complicated. However, the best recognition accuracies of event three and event two were generally higher than that of event one. The recognition accuracies of event three and event two were 96.16% and 93%, respectively. We can conclude that the driving scenario and the number of input features have considerable impacts on the recognition accuracy of the algorithm.

Compared with the traditional SVM algorithm model, the recognition accuracy of the Bi-LSTM model with

attention mechanism in the total event reached 94.33%, which is a high accuracy of distraction recognition. This result is 3.83% higher than that of the SVM model. The model proposed in this paper indicates reasonable applicability for distraction recognition in mixed traffic environments, which can be applied in an autonomous driving assistance system. It is of practical significance for the driver to effectively avoid road traffic accidents and improve driving efficiency in mixed traffic environments.

Limitations of this study are as follows:

- (i) The Bi-LSTM model with attention mechanism discussed in this paper does not perform real-time recognition
- (ii) There is a certain gap between the created mixed traffic environment and a real mixed traffic environment

## 5. Conclusion

This paper proposes a method based on Bi-LSTM with attention mechanism to identify the cognitive distraction state of a driver. An on-road experiment was implemented at a nonsignalized mixed traffic environment intersection, in which the driver's physiological state information, driving performance data, and eye movement data were collected. The algorithm uses the optimal feature combinations of driving performance-related features and eye movement features as inputs to identify distracted driving, and the algorithm is compared with the traditional SVM model. The following conclusions can be summarized as follows: (1) The weight value of the eye movement features is greater than the weight value of the driving performance-related features in the first-level evaluation indices. This result indicates that the eye movement features are very useful for identifying cognitive distraction in mixed traffic environments. (2) The cognitive load required for drivers in mixed traffic scenarios is higher than that required in traditional traffic scenarios, and the higher is the likelihood of distracted driving. (3) The cognitive distraction state recognition of drivers based on Bi-LSTM with attention mechanism provided more advantages in mixed traffic environments, which is 3.83% higher than the SVM model in the total event. This model provides a foundation for intelligent in-vehicle devices to detect driver distraction, which can reduce the number of accidents caused by distracted driving.

In this study, the cognitive distraction recognition model does not perform real-time recognition; therefore, future research can improve this model and transform it into a real-time recognition model.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Qiang Hua wrote the main manuscript text. Lisheng Jin established the experimental platform and provided financial support. Yuying Jiang conducted the experiments. Ming Gao and Baicang Guo improved the manuscript.

## Acknowledgments

This work was supported by the National Key R&D Program of China under Grant 2018YFB1600501, the National Natural Science Foundation of China (nos. U19A2069 and 52072333); the National 13th Five-Year Plan's Science and Technology Project of the Education Department of Jilin Province (JJKH20200988KJ), the Natural Science Foundation of Hebei Province (no. E2020203092), and Hebei Provincial Key Research Projects (no. 20310801D). The contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors are acknowledged.

## References

- [1] D. Omid, R. Vikas, and T. Mojtaba, "Wearable driver distraction identification on-the-road via continuous decomposition of galvanic skin responses," *Sensors*, vol. 18, no. 2, pp. 503–518, 2018.
- [2] National Highway Traffic Safety Administration, *Distracted Driving*, <https://www.nhtsa.gov/risky-driving/distracted-driving>, 2017.
- [3] S. P. Mcevoy, M. R. Stevenson, and M. Woodward, "The prevalence of, and factors associated with, serious crashes involving a distracting activity," *Accident Analysis & Prevention*, vol. 39, no. 3, pp. 475–482, 2007.
- [4] G. Li, W. Lai, X. Sui et al., "Influence of traffic congestion on driver behavior in post-congestion driving," *Accident Analysis & Prevention*, vol. 141, p. 1055508, 2020.
- [5] J. Pérez, F. Seco, V. Milanés, A. Jiménez, J. C. Díaz, and T. De Pedro, "An RFID-based intelligent vehicle speed controller using active traffic signals," *Sensors*, vol. 10, no. 6, pp. 5872–5887, 2010.
- [6] C. B. Néstor, A. S. Raúl, M. E. Pedro, A.-V. José, E.-B. Arthur, and M. C. Aldo, "Traffic congestion detection system through connected vehicles and big data," *Sensors*, vol. 16, no. 5, pp. 599–625, 2016.
- [7] D. Isele, R. Rahimi, A. Cosgun, K. Subramanian, and K. Fujimura, "Navigating occluded intersections with autonomous vehicles using deep reinforcement learning," in *Proceedings of the International Conference on Robotics and Automation*, pp. 2034–2039, Brisbane, Australia, May 2018.
- [8] B. K. Pathivada and V. Perumal, "Analyzing dilemma driver behavior at signalized intersection under mixed traffic conditions," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 111–120, 2019.
- [9] Y. Liang and J. D. Lee, "Combining cognitive and visual distraction: less than the sum of its parts," *Accident Analysis and Prevention*, vol. 42, no. 3, pp. 881–890, 2010.
- [10] L. Li, B. Zhong, C. J. Hutmacher et al., "Detection of driver manual distraction via image-based hand and ear recognition," *Accident; Analysis and Prevention*, vol. 137, 2020.
- [11] T. C. Ojsteršek and D. Topolšek, "Influence of drivers' visual and cognitive attention on their perception of changes in the traffic environment," *European Transport Research Review*, vol. 11, no. 1, pp. 1–9, 2019.

- [12] A. Olia, S. Razavi, B. Abdulhai, and H. Abdelgawad, "Traffic capacity implications of automated vehicles mixed with regular vehicles," *Journal of Intelligent Transportation Systems*, vol. 22, no. 3, pp. 244–262, 2018.
- [13] J. Wang, R. Knipling, and M. Goodman, "The role of driver inattention in crashes: new statistics from the 1995 crash worthiness data system (CDS)," in *Proceedings of the 40th Annual Proceedings: Association for the Advancement of Automotive Medicine*, pp. 377–392, Vancouver, Canada, October 1996.
- [14] P. Choudhary and N. R. Velaga, "Gap acceptance behavior at unsignalized intersections: effects of using a phone and a music player while driving," *Traffic Injury Prevention*, vol. 20, no. 4, pp. 372–377, 2019.
- [15] G. Li, Y. Wang, F. Zhu et al., "Drivers' visual scanning behavior at signalized and unsignalized intersections: a naturalistic driving study in China," *Journal of Safety Research*, vol. 71, pp. 219–229, 2019.
- [16] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 340–350, 2007.
- [17] M. Atiquzzaman, Y. Qi, and R. Fries, "Real-time detection of drivers' texting and eating behavior based on vehicle dynamics," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 58, no. 10, pp. 594–604, 2018.
- [18] D. W. Hosmer and S. Lemeshow, "Goodness of fit tests for the multiple logistic regression model," *Communications in Statistics - Theory and Methods*, vol. 9, no. 10, pp. 1043–1069, 1980.
- [19] H. Zhang, D. L. Qian, and C. F. Shao, "Discrimination of driver distraction in simulated driving environment," *Journal of China Highway and Transport*, vol. 31, no. 4, pp. 43–51, 2018.
- [20] Y. Liao, S. E. Li, W. Wang, Y. Wang, G. Li, and B. Cheng, "Detection of driver cognitive distraction: a comparison study of stop-controlled intersection and speed-limited highway," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1628–1637, 2016.
- [21] C. Craye and F. Karray, "Driver distraction detection and recognition using RGB-D sensor," *Computer Science*, <http://arxiv.org/abs/1502.00250>, 2015.
- [22] J. Son and M. Park, "Detection of cognitive and visual distraction using radial basis probabilistic neural networks," *International Journal of Automotive Technology*, vol. 19, no. 5, pp. 935–940, 2018.
- [23] H. Zhang, Y. Liu, C. Wang, R. Fu, Q. Sun, and Z. Li, "Research on a pedestrian crossing intention recognition model based on natural observation data," *Sensors*, vol. 20, no. 6, pp. 1776–1797, 2020.
- [24] W. Wang, S. Yang, Y. Jiang, S. Zhang, and W. Deng, "What contributes to driving behavior prediction at unsignalized intersections?," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 100–114, 2019.
- [25] J. Wu and H. Xu, "Driver behavior analysis for right-turn drivers at signalized intersections using shrp 2 naturalistic driving study data," *Journal of Safety Research*, vol. 63, pp. 177–185, 2017.
- [26] C. Cabrall, N. Janssen, J. Goncalves, A. Morando, M. Sassman, and J. D. Winter, "Eye-based driver state monitor of distraction, drowsiness, and cognitive load for transitions of control in automated driving," in *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1981–1982, Budapest, Hungary, October 2016.
- [27] S. W. Savage, D. D. Potter, and B. W. Tatler, "Does preoccupation impair hazard perception? a simultaneous EEG and Eye Tracking study," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 17, no. 2, pp. 52–62, 2013.
- [28] Y. Liao, S. E. Li, W. Wang, Y. Wang, G. Li, and B. Cheng, "The impact of driver cognitive distraction on vehicle performance at stop-controlled intersections," in *Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV)*, pp. 946–951, Seoul, Korea (South), June 2015.
- [29] D. L. Strayer and F. A. Drews, "Cell-Phone-induced driver distraction," *Current Directions in Psychological Science*, vol. 16, no. 3, pp. 128–131, 2007.
- [30] M. J. Kane, A. R. A. Conway, T. K. Miura, and G. J. H. Colflesh, "Working memory, attention control, and the N-back task: a question of construct validity," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 33, no. 3, pp. 615–622, 2007.
- [31] V. Alizadeh and O. R. Dehzangi, "The impact of secondary tasks on drivers during naturalistic driving: analysis of EEG dynamics," in *Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2493–2499, Rio de Janeiro, Brazil, November 2016.
- [32] Y. L. Pei, Y. Q. Jin, and H. F. Chen, "Experimental study on EEG characteristics of driving fatigue in different ages," *Journal of China Highway and Transport*, vol. 31, no. 4, pp. 59–65, 2018.
- [33] J. Lee, B. D. Sawyer, B. Mehler et al., "Linking the detection response task and the attend algorithm through assessment of human-machine interface workload," *Transportation research record: Journal of the Transportation Research Board*, vol. 2663, pp. 82–89, 2017.
- [34] L. S. Jin, B. C. Guo, X. Y. Jiang, F. Wang, X. Xie, and M. Gao, "Study on the impact degrees of several driving behaviors when driving while performing secondary tasks," *IEEE Access*, vol. 6, p. 1, 2018.
- [35] J. Li, Z. Q. Yue, and S. C. Wong, "Performance evaluation of signalized urban intersections under mixed traffic conditions by gray system theory," *Journal of Transportation Engineering*, vol. 130, no. 1, pp. 113–121, 2004.
- [36] F. R. Wang, B. C. Guo, L. S. Jin, L. L. Gao, and X. Y. Yue, "Selection of safe evaluation indexes and weight calculation for secondary task driving," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 47, no. 6, pp. 1710–1715, 2017.
- [37] Z. Wu, "The application of FAHP in decisions of pavement maintenance," in *Proceedings of the 3rd International Conference on Energy Materials and Environment Engineering*, pp. 1755–1315, Bristol, England, April 2017.
- [38] G. C. Cawley and N. L. C. Talbot, "Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers," *Pattern Recognition*, vol. 36, no. 11, pp. 2585–2592, 2003.
- [39] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [40] C. C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems & Technology*, vol. 2, no. 3, p. 27, 2011.
- [41] J. Engström, G. Markkula, T. Victor, and N. Merat, "Effects of cognitive load on driving performance: the cognitive control hypothesis," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 59, no. 5, pp. 734–764, 2017.