

## Research Article

# Identification of Driver Distraction Based on SHRP2 Naturalistic Driving Study

Zhiqiang Liu , Shiheng Ren , and Mancai Peng 

*School of Automobile and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China*

Correspondence should be addressed to Shiheng Ren; [renshiheng163@163.com](mailto:renshiheng163@163.com)

Received 18 December 2020; Revised 10 April 2021; Accepted 22 April 2021; Published 4 May 2021

Academic Editor: Ali Ahmadian

Copyright © 2021 Zhiqiang Liu 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.

Abundant evidence shows that driver distraction is one of the fundamental causes of traffic accidents. Current detected methods of driver distraction are mostly based on intrusive or semi-intrusive. The methods not only interfere with the driving task but also are restricted by various environmental factors, resulting in a high false positive rate. This paper only considers noninvasive vehicle kinematics indicators and proposes a recognition method based on deep learning. Firstly, some car following segments are obtained from the naturalistic driving database, and typical distracted segments are extracted by using situation awareness. Then, distracted recognition indexes' set is established and only contains vehicle kinematic features. Thirdly, the gradient boosted decision tree recursive feature elimination (GBDT-RFE) and random forest recursive feature elimination (RF-RFE) are used to rank the importance of features. The indexes with higher importance are obtained. Finally, the long short-term memory neural network (LSTM-NN) is utilized to realize the classification and recognition of distracted driving, and the results are compared with SVM and AdaBoost. The results show that the F1-scores of LSTM-NN are 89% and 91% in distracted and normal driving, which are higher than SVM and AdaBoost. The average F1-score of distracted recognition (12% and 7%) is higher than SVM and AdaBoost. The false positive rate of different distracted types is less than 15%. LSTM-NN can effectively learn the information before and after the distracted sequence, which is conducive to accurately estimate the driver's attention state. The study provides a method for vehicle distraction warning system and driving risk propensity assessment.

## 1. Introduction

Driver distraction is one of the main reasons for frequent traffic accidents. According to the national highway traffic safety administration (NHTSA), 25%–30% of accidents in the United States are related to driving distraction [1]. With the development of intelligent network connection, the on-board information and service system have brought convenience to communication and interaction. But it also distracts drivers at the same time, which leads to traffic safety problems. Therefore, analysis and recognition of distracted driving behavior are of great significance for researchers to build a real-time distracted driver monitoring system and reduce the number of traffic accidents [2, 3].

Scholars have put forward many definitions of distraction. Regan and Hallett proposed that driver's distraction can be regarded as a form of inattention in driving [4]. It is mainly

performed when driver's attention is shifted from key activities of safety driving to the activities which are competitive with driving task. The behavior leads to insufficient attention to key activities. According to this definition, distracted driving can be divided into visual, cognitive, and operational distraction. Visual distraction refers to driver's attention leaving the road for a short time. For example, the driver may adjust the car stereos or check the mobile phone message during driving, which has a significant impact on lane keeping, steering rate controlling, and braking reaction time [5, 6]. Cognitive distraction refers to driver's thought being disordered, which affects driving performance by reducing the attention allocated to driving scene and processing pre-determined information [7–9]. It also reduces driver's ability to detect targets in the whole driving environment [10]. Operational distraction, such as adjusting the interior sound and sending text messages [11], requires the driver's hand to

leave the position related to the driving task and operate other things [12, 13]. The above three forms rarely exist separately; more often two or more forms occur simultaneously, such as visual-operational distraction [14] and visual-cognitive distraction [15]. Current studies usually set driving subtask for a certain form of distraction to obtain driver's distracted segments [16]. However, distraction is a product of complex interaction among driver, vehicle, and road environment. The performance is not fixed and it occurs randomly. Therefore, it does not make much sense to study a certain type of distraction separately.

Nowadays, there are two main methods (driving simulator test and real car experiment test) for distracted driving behavior. The real vehicle test can be subdivided into field-controlled experiment and naturalistic driving experiment [17, 18]. A driving simulator has the advantages of low cost and convenient operation [19], but drivers' behaviors are not natural enough and the risk of operation awareness is lower. Compared with that, the data authenticity of real car test is higher. However, it is time-consuming and requires many types of equipment. This method belongs to controlled experiment, which reduces the interaction between the driver and environment. The test cannot reflect distraction under natural driving in reality [20]. In recent years, many researches on natural driving provide data support for the analysis of driving behaviors in real environments [21–23]. However, few studies obtain and analyze distracted databases from uncontrolled natural driving experiments. On the other hand, the data types of distracted research can be divided into three categories. The first type is based on driving performance characteristics [24], such as eye movement [25, 26], psychosomatic response [27], and head posture [28, 29]. But detection method based on physiological signal is limited by the heavy equipment, high cost, and intrusive detection to the driver. So it is difficult to be a large-scale application in practice. Although the detection based on eye movement or head posture is widely used in researches, it is greatly affected by external environment. For example, light changing and vehicle vibration result in a high false positive rate. The second type uses the data captured by various sensors during vehicle driving, such as controller area network (CAN) bus and radar, including speed, acceleration, and relative distance. This kind of researches directly recognizes driver's attention by observing vehicle's kinematics response. The data source is simple and the kinds of vehicle kinematic features are various, which helps to achieve higher classification accuracy [30]. The third type is data fusion, which has more comprehensive features. However, the relationship between distracted driving behavior and vehicle kinematics characteristics is not clear. Some usual behaviors, such as driving with one hand and simple communication with passengers, do not affect driver's situation awareness ability. And vehicle kinematics characteristics also do not have abnormal changes. But this behavior leads to confusion of distracted recognition after mixed data [31].

The advantages of vehicle kinematics features are a great variety and easy to obtain, so it is widely adopted in related papers. According to the range of features, it can be divided into lateral control features (lane offset, lateral acceleration, etc.), longitudinal control features (speed, acceleration, etc.),

and car following features (the relative distance, speed with the front vehicle, etc.) [32]. Kountouriotis et al. studied the effect of cognitive distraction on the steering wheel rate (SR) [33]. Owing to various vehicle kinematics features, the redundant characteristics may affect each other. They could reduce the computational efficiency and affect the accuracy of the model. Therefore, the features' subset needs to be screened. Li et al., based on a natural driving experiment, utilized statistical analysis to select two indicators with obvious differences between distraction and normal driving (ND) and then adopted the recognition model to identify the driver status [14].

With the deepening of machine learning researches, different classification methods are widely adopted in the field of distracted driving. The core of classification is to uncover the hidden associated information between indicators and distraction. Some scholars make use of machine learning to aggregate the distracting time-series segments by mean, variance, and extreme values. Such models include support vector machine (SVM) [17], AdaBoost [34], and Bayesian network [35, 36]. However, distraction is a dynamic multi-interaction process and occurs randomly. The duration and influence degree are also different. Therefore, this kind of method cannot make accurate identification of driving state by using time series. On the contrary, some methods considering time sequence can explore the dynamic and multi-interactive process. Liang and Lee found that distraction was highly time-dependent when predicting the current drivers' state [37]. Therefore, a model based on a dynamic Bayesian network was developed to detect cognitive distraction using behavioral parameters of eyes and drivers. With the development of deep learning technology, more and more researches make use of different neural networks (NN), especially recurrent neural networks (RNN), to deal with time-series problems. Wollmer et al. introduced an online driver distraction detection method, which collects data through natural driving experiment [38]. Then they used long short-term memory neural network (LSTM-NN) to realize the detection of distraction. However, all of these studies' data are collected under controlled driving test or simulated driving test, which have quite limitations.

Although previous studies have achieved some successful cases in the recognition of driver distraction, few works are dedicated to the use of natural driving data for driver distraction recognition. In view of that, this paper proposes a distraction identification method based on deep learning. The distracted segments are extracted from completely natural driving data. Then, the vehicle kinematics features' set is established. After that, the LSTM-NN model is set up to recognize binary classification and multiclassification. The effectiveness of the model is verified by comparing with the classical SVM and AdaBoost. The contribution of this paper mainly includes the following points.

This paper proposes a method based on vehicle kinematics situation to obtain distracted events. The data effectively supports the acquisition in completely naturalistic driving. This paper utilizes machine learning to select features with high importance, which reduces the input dimensions of classification and improves the output accuracy

of the model. This paper develops a model based on the time series of LSTM-NN, which can solve the problem of time dependence and achieve higher accuracy compared with other models. It also has a high ability when recognizing multiclassification distracted situations. Moreover, it provides a new way for recognition the distracted driving based on abnormal vehicle kinematics characteristics.

The paper is organized as follows. Section 2 discusses the method we used in this paper. Section 3 gives an overview of the proposed approach and introduces a recognition model of distracted state based on LSTM-NN. Section 4 illustrates the application of LSTM-NN to the identification of driver distraction and the comparisons with SVM and AdaBoost. Finally, in Section 5, we draw the conclusions and present possible research directions.

## 2. Methods

**2.1. SHRP2 Database.** This paper relies on a sampling of data from SHRP 2 NDS, in which participants' vehicles were equipped with four video cameras (forward, driver, rear, and passenger snapshot view) over a period of up to three years. Within this study, 35 million vehicle miles and 5.4 million trips from approximately 3500 U.S. drivers were recorded. A complete description of SHRP 2 data is located on the study's website (available at <https://insight.SHRP2nds.us/>). The frequency range of data collected by the acquisition system is 10 to 50 Hz, including position information collected from GPS, acceleration of gyroscope, and vehicle kinematics characteristics of the CAN bus. The relative speed and distance of the surrounding vehicles are detected by radar, and video of four cameras as shown in Figure 1. Previous studies have shown that drivers of different ages have no significant difference in distracted behavior [39]. Therefore, this study selected 15 drivers, including 12 men and 3 women. All the drivers are physically and mentally healthy without any physical dysfunction. The driving mileage is more than 5000 km, and the age is between 33 and 45 years. Their driving experience is basically the same.

**2.2. Car Following Sample Extraction.** The original data contains a series of driving scenarios, such as parking and waiting, lane changing and overtaking, and stable car following. However, in addition to stable car following, other behaviors have specific requirements for drivers' psychological load. So the possibility of distraction is small for this behavior. To analyze the driver's distracted behavior more accurately, this paper extracts segments of the stable car following from the natural driving database. Due to the existence of noise interference and other problems, it is necessary to preprocess the data before using. The 3-sigma principle is adopted to eliminate the outliers. Then the missing values are filled with linear interpolation to ensure that the sampling frequency of all car following segments is 10 Hz. The car following segments are obtained from the original data by setting conditions. In this paper, they are defined as follows.

(1) The identification number of radar targets remains unchanged, which indicates the front vehicle has not changed. (2) The lateral offset of adjacent timestamps has no abrupt change greater than 1.5 m, which ensures that the vehicle does not change lanes or overtake. (3) The longitudinal distance from the front vehicle is less than 100 m, which eliminates the completely unrestricted car following state. (4) When the speed of the test vehicle is greater than 20 km/h, waiting event is excluded with low speed. (5) Car following time is equal to 10 s, which ensures that each car following segment contains enough sequences for subsequent method.

According to the rules, 36734 car following segments were extracted. After comparison of sampling, 34080 effective samples were found, accounting for 92.8% of the total number. The total length of effective segments was about 81 hours.

**2.3. Index of Vehicle Kinematics.** Referring to the previous researches, the data utilized in this study come from the vehicle kinematics data of CAN bus and external radar. The data can truly reflect abnormal vehicle kinematics caused by distraction. These basic features can be divided into three categories (lateral control, longitudinal control, and car following) according to the vehicle kinematics characteristics.

Steering wheel rate entropy (SRE) is calculated from SR in the original data, which reflects the steering smoothness. A dedicated driver can constantly assess the environmental situation and correct steering control unconsciously. But distracted drivers tend to take large actions to correct errors, resulting in lower smoothness [40]. This paper mainly focuses on driving characteristics which are directly related to driving control ability. The features directly quantify the extent to which driving behavior is affected, rather than how the driver's physiological or mental state is affected. The SR prediction is shown in Figure 2. At time  $t$ , the original SR is  $\theta(t)$ , and the predicted SR value  $\theta_p(t)$  is obtained by the second-order Taylor expansion. Then the difference between the actual value and the predicted value is calculated to obtain  $e(t)$ , as shown in formula (1). Considering the sampling frequency, 1s is selected as the prediction period. Finally, the SRE is obtained by formula (1). And  $f_i$  is the frequency of the error falling into the  $i$ -th interval. The SRE is higher, the steering stability is worse, and the driving behavior is affected to a great extent.

$$\begin{aligned} \theta_p(t) &= \theta(t-1) + (\theta(t-1) - \theta(t-2)) \\ &\quad + \frac{1}{2} (\theta(t-1) - \theta(t-2)) - (\theta(t-2) - \theta(t-3)), \\ e(t) &= \theta(t) - \theta_p(t), \\ SRE &= - \sum_{i=1}^9 f_i \log_9(f_i). \end{aligned} \tag{1}$$

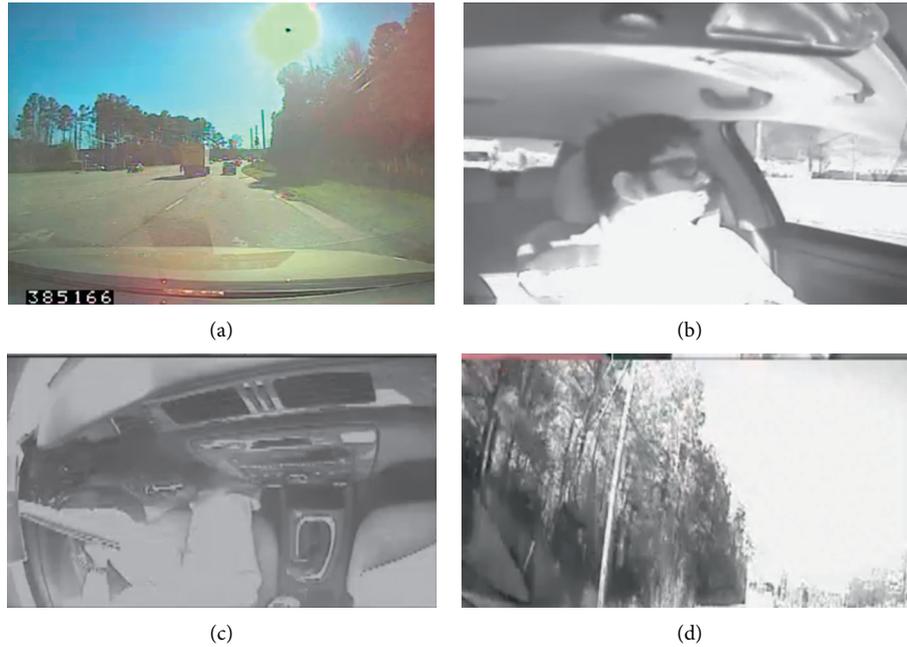


FIGURE 1: Four-camera views.

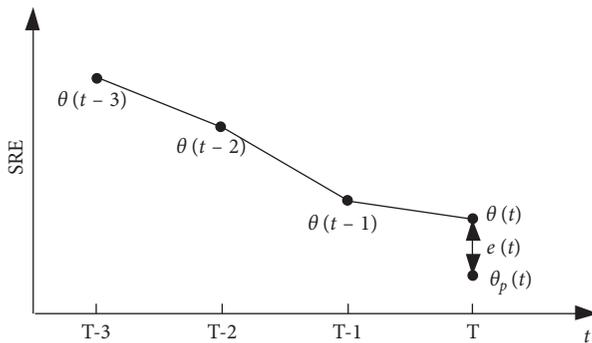


FIGURE 2: Prediction value and relative error of SRE.

The desired reaction time (DRT) refers to the potential reserved reaction time for handling emergencies when the driver follows the front vehicle. The scenario diagram is shown in Figure 3. This paper supposes that  $b_f$  and  $b_e$  are the acceleration of front and rear vehicles and they are  $5.8 \text{ m/s}^2$  at any time.  $V_f$  and  $V_e$  represent the speed of the front and rear vehicles. If the front car brakes at  $t_1$  and stops at  $t_3$ , the rear car takes emergency braking at  $t_2$  after the reaction time  $\tau_e$  (set as 1s in this paper). DRT describes the time required for the change of distance between the rear and the front. In the process of car following, DRT is related to the driver's personal characteristics. The DRT is smaller, and driver's reaction time is shorter to deal with the unexpected situation. And his driving style is more aggressive. Compared with the common time to collision [41], DRT can describe the safety of car following better when the speed of the rear vehicle is less than that of the front.

In order to deeply discuss vehicle kinematics features, this paper extends the basic features into six descriptive statistical indicators, which are mean, standard deviation,

maximum, minimum, residual sum of squares, and regression slope. Each basic characteristic is expanded into six indicators except SRE. For example, lateral acceleration is expanded into mean of lateral acceleration, standard deviation of lateral acceleration, maximum of lateral acceleration, minimum of lateral acceleration, residual sum squares of lateral acceleration, and regression slope of lateral acceleration. The descriptive statistical indexes are calculated, and finally, 43 vehicle kinematics characteristics are obtained as shown in Table 1.

**2.4. Distracted Segments Extraction Based on Situation Awareness.** Situation awareness is defined as the requirements of what does this event mean to the operator and what might be happening next [42]. In terms of driving performance, situation awareness relies on the driving perception of road elements, cognition of current driving tasks, and prediction of surrounding vehicles. Ultimately, drivers change or adjust current driving operations after processing this information. Effective situation awareness is essential for driving safety, including monitoring and updating the position of other vehicles, road conditions, and proper speed [15]. Relevant researches show that there is a significant correlation between situation awareness ability and driving performance during distraction. For example, when a driver makes a call during driving, the behavior weakens his ability of perception and prediction. It also increases the reaction time, decreases the control ability, and loses the initiative to make a decision. According to cognitive behavior theory [43], the change of driver's attention will be reflected in the change of vehicle dynamic features, which provides a prerequisite for extracting distracted segments. Combined with the macro driving performance of vehicle control under

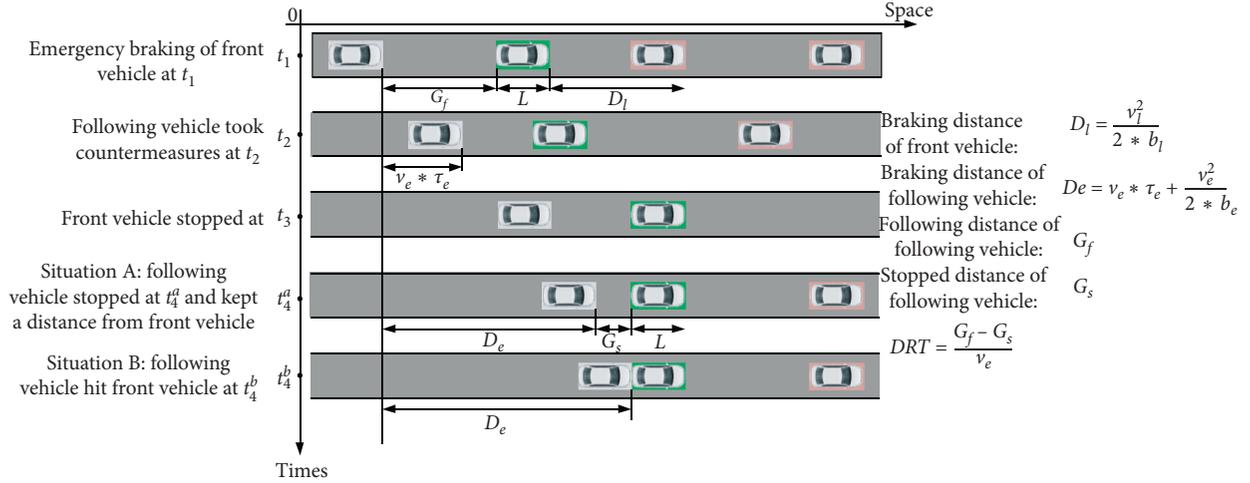


FIGURE 3: Scenario of desired reaction time (DRT).

TABLE 1: Vehicle kinematic features.

Statistical indicators	Characteristic	Unit	Signal
Mean	Lateral acceleration	$m \cdot s^{-2}$	Accel_y
Standard deviation	Steering wheel rate	$rad \cdot s^{-1}$	SWR
Maximum	Longitudinal acceleration	$m \cdot s^{-2}$	Accel_x
Minimum	Desired reaction time (DRT)	s	DRT
Residuals sum of squares	Relative velocity	$m \cdot s^{-1}$	Rel_V
Slope of regression	Distance of headway Velocity	m $m \cdot s^{-1}$	DHW V
Mean	SRE		SRE

distraction, three typical scenarios and their identification conditions are proposed in this paper.

- (i) Following distance larger (FDL): the relative distance between the rear and front vehicle is getting larger under no special situations.

Description: when the front vehicle does not decelerate (lasting up to  $T_1$ ), if the DRT and the relative distance increase, the speed of the rear vehicle does not increase. The segment is considered to belong to FDL.

- (ii) Following distance smaller (FDS): delayed braking results in a small distance from the front vehicle.

Description: when the front vehicle decelerates (lasting up to  $T_2$ ), the DRT decreases below the danger threshold, and the speed of the rear vehicle does not decrease significantly. It is considered that the segment belongs to FDS.

- (iii) Following distance unsteadily (FDU): the vehicle swings unsteadily.

Description: extreme values of lateral acceleration or SR exceed the threshold range.

Due to the influence of road smoothness and driver's microadjustment, lateral acceleration, SR and acceleration

fluctuate continuously in the normal range. It is not conducive to the identification of vehicle kinematics trend. Therefore, adaptive Kalman filter is adopted to smooth the correlation feature sequence so as to obtain a stable trend. In order to eliminate the differences in driving styles and habits of drivers, this paper establishes a historical data set of indicators for each driver. Then we set different thresholds according to the cumulative curve. Considering the complexity of the road environment, this research selects as many effective distraction scenarios as possible. This paper appropriately reduces the minimum threshold of DRT and selects the 35 percentile of cumulative distribution as the condition of FDS. 90 and 90 percentiles are adopted as the lateral acceleration and SR as the extraction threshold of FDU for each driver. The cumulative curves of the three indicators are shown in Figure 4. These thresholds will be used as an important reference for distinguishing different drivers when they take emergency measures under the same subsituation.  $g$  is the acceleration of gravity with a score of  $9.8 m/s^2$ . It is worthy to point out that some segments contain more than one distracted scenario. In order to deal with this problem, we check all the segments to search whether there are distracted scenarios (FDL, FDS, and FDU) and calculate the time that belongs to each scene. The scene with the largest proportion of duration was regarded as the main type of distraction. According to the above conditions, a total of 3133 suspected distracted segments were obtained.

However, there may be wrongly identified in the distracted segments only by situation awareness. Actually, it can be explained by the abnormal data and driver's personality behavior, such as driving aggressively and being extremely cautious. In addition, there must be visual and manual distraction when the driver is operating the on-board equipment or talking with the passenger frequently. Pure cognitive distraction or inattention is excluded because it is hard to monitor and detect. In order to filter out the above events, we organized twelve postgraduates in a group of three to manually check the suspected segments by using the front and hand videos. According to the sensor data, we can determine whether each suspected segment is caused by

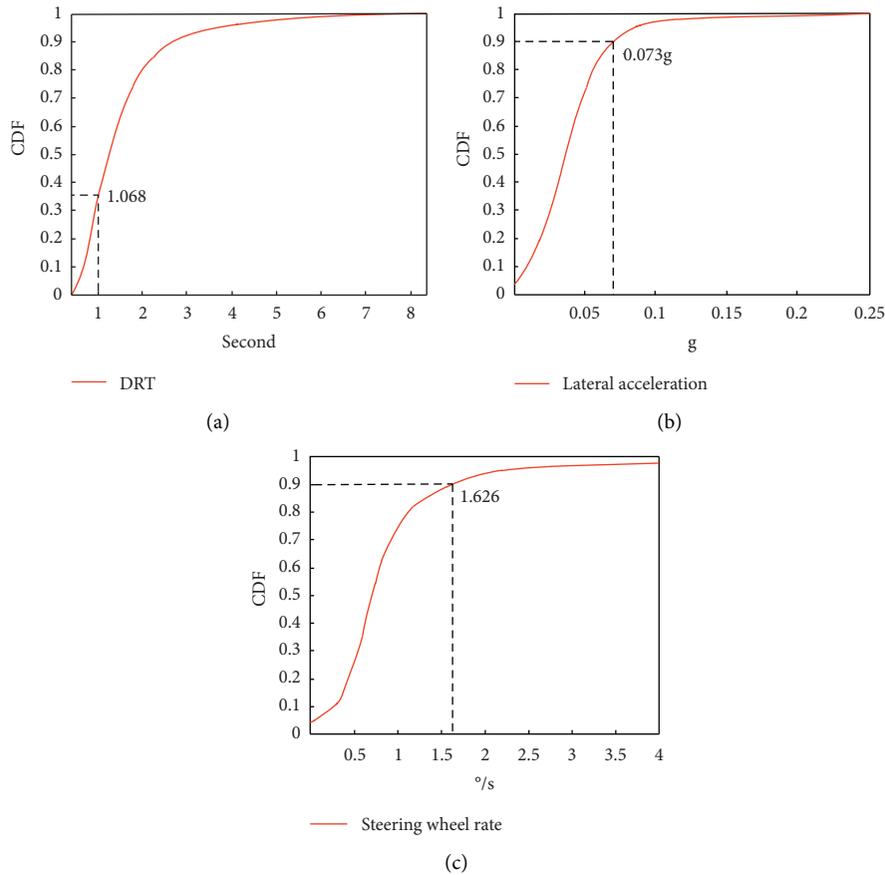


FIGURE 4: Cumulative probability distribution curve. (a) Cumulative curve of DRT. (b) Cumulative curve of lateral acceleration. (c) Cumulative curve of steering wheel rate.

the driver's obvious distracted behavior. All members of each group reviewed the same data set. If the driver's behavior was not distracted after cross-validation, the segment was regarded as none distraction and deleted from the database. Otherwise, the segment was retained as a distracted sample. The normal samples are randomly selected from the car following database. After manually checking the videos and cross-validation, obvious different from distraction behavior can be regarded as be normal. Through the above steps, 2400 normal segments were obtained, which were labeled as ND. 596 segments were labeled as FDL. 324 segments were labeled as FDS. 277 segments were labeled as FDU. FDL, FDS, and FDU are all called distracted segments, and there are 1197 segments in total.

**2.5. Selection Indexes of Distracted Driving.** In this study, a total of 43 indicators are selected as the primary indexes set. In order to reduce the interference of redundant indexes and keep the physical meaning of indexes, feature ranking of importance is used to achieve dimension reduction. Firstly, the min-max normalization is adopted to reduce the differences between different scale features. Then, the Pearson correlation coefficient is used to test the linear correlation between features. If the features whose absolute score of the correlation coefficient is greater than 0.8, it will be deleted. It

can be found that the average value of basic features is highly correlated with its maximum and minimum. For example, the average DRT is highly positively correlated with its standard deviation, maximum, minimum, and residual sum of squares. In addition, there is a significant correlation between some indicators which belong to the same vehicle dynamic feature. For example, the correlation coefficient of lateral acceleration regression slope and SR regression slope is 0.86. And the correlation coefficient between relative speed and minimum score of headway is 0.89. With [44], the mean that is lower affected by outliers is retained in the significant correlation indexes. Finally, thirty indexes are obtained.

Two integrated learning methods are utilized to obtain important indicators. The first method is gradient boosting decision tree recursive feature elimination (GBDT-RFE). GBDT is a common boosting algorithm. It can adjust the distribution of training samples according to the performance of based learners. So the training samples, which were classified wrongly by previous based learners, could get more attention in the follow-up steps. Finally, based learners are weighted and combined to obtain the optimal classifier [45]. The second method is random forest recursive feature elimination (RF-RFE). RF is one of the bagging methods. On the basis of decision tree integration, random attribute selection is further selected in the training process. The output index categories are determined by the vote of each tree.

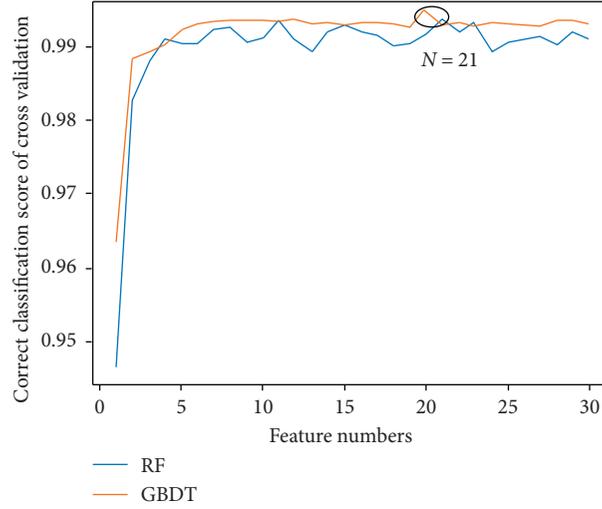


FIGURE 5: The relationship between feature numbers and classification score.

GBDT-RFE and RF-RFE, respectively, use GBDT and RF to analyze the importance of indicators. After getting the ranking, features with higher importance through RFE are selected. 5-fold cross-validation is adopted in this paper. The relationship between feature numbers of two methods and the corrected classification score of cross-validation is shown in Figure 5. It can be found that when the number of features  $N$  equal to 21, the classification accuracy of GBDT-RFE and RF-RFE is close to the optimal score at the same time. The coincidence rate of the top 21 indexes obtained by the two methods is more than 80%, and the ranking results of importance were basically the same. Therefore, GBDT-RFE with a higher average correct classification score was selected to obtain the first 21 indicators, as shown in Table 2.

### 3. Identification Model of Distracted State

**3.1. LSTM-NN Model Overview.** Distraction is a continuous process with dynamic behavior fluctuations over time. Most of the classification methods adopted in previous studies focused on the features reflecting the whole sequence, but it was difficult to capture the change of distracted events process. Considering the time dependence of distracted events, driver's attention can be identified by time-series information. In recent years, deep learning has shown great potential in dealing with multitime-series classification. RNN is a kind of artificial neural network in which nodes are connected in a ring directionally. The internal state of this network can show drivers' dynamic behavior. Different from feedforward neural network, RNN can use internal memory to process sequence data, which makes it often utilized to deal with some complex deep learning problems. However, due to the gradient vanishing effect, the simple RNN cannot deal with the long sequence dependence well. In order to overcome this limitation, Hochreiter and Schmidhuber proposed LSTM-NN, as shown in Figure 6 [46]. Each hidden unit in LSTM-NN contains one or more memory cells. The operation of these memory cells is controlled by the internal input gate  $i_t$ , forgetting gate  $f_t$ , and output gate  $o_t$ . The input gate determines which part of the

input value can be used to update the memory state. The forgetting gate determines the information to be retained or removed in the memory unit. And the output gate determines the output content. The expression of the three gates at time  $t$  can be expressed by the following formula:

(i) Input gate:

$$i_t = \sigma[W_i(x_t + h_{t-1} + C_{t-1}) + b_i]. \quad (2)$$

(ii) Forget gate:

$$\begin{aligned} f_t &= \sigma[W_f(x_t + h_{t-1} + C_{t-1}) + b_f], \\ \bar{C}_t &= \tan h[W_c(x_t + h_{t-1}) + b_c]. \end{aligned} \quad (3)$$

(iii) Memory cell unit:

$$C_t = f_t \times C_{t-1} + i_t \times \bar{C}_t. \quad (4)$$

(iv) Output gate:

$$\begin{aligned} o_t &= \sigma[W_o(x_t + h_{t-1} + C_{t-1}) + b_o], \\ h_t &= o_t \times \tan h(C_t), \end{aligned} \quad (5)$$

where  $i_t$ ,  $f_t$ , and  $o_t$  are the score of input gate, forgetting gate, and output gate after passing through activation function at time  $t$ .  $W_i$ ,  $W_f$ ,  $W_c$ ,  $W_o$ , and  $W_f$  and  $b_i$ ,  $b_f$ ,  $b_c$ ,  $b_o$ , and  $b_f$  are weight and deviation.  $C_t$  is an alternative value of memory cell state  $C_t$  at time  $t$ .  $x_t$  and  $h_t$  are the input of the memory cell and the output of the final decision. At each time step, the LSTM-NN can read, write, or reset the memory unit through the three gates. This strategy allows LSTM-NN to memorize and access the information before multistep.

TABLE 2: Ranking table of important feature indicators.

Importance ranking	Characteristic	Signal	Importance ranking	Characteristic	Signal
1	Mean of velocity	V_mean	12	Residuals sum of squares of velocity	V_ess
2	Mean of desired reaction time	DRT_mean	13	Mean of longitudinal acceleration	Accel_x_mean
3	Mean of headway distance	HWD_mean	14	Regression of slope of desired reaction time	DRT_sr
4	Maximum of relative velocity	Rel_V_mean	15	Standard deviation of longitudinal acceleration	Accel_x_std
5	Standard deviation of steering wheel rate	SWR_std	16	Regression of slope of relative velocity	Rel_V_sr
6	Regression of slope of lateral acceleration	Accel_y_sr	17	Mean of steering wheel rate	SWR_mean
7	Regression of slope of velocity	V_sr	18	Mean of lateral acceleration	Accel_y_mean
8	Steering rate entropy	SRE	19	Regression of slope of longitudinal acceleration	Accel_x_sr
9	Standard deviation of headway distance	HWD_std	20	Residuals sum of squares of steering wheel rate	SWR_ess
10	Residuals sum of squares of headway distance	HWD_ess	21	Residuals sum of squares of relative velocity	Rel_V_ess
11	Maximum of lateral acceleration	Accl_y_max	—	—	—

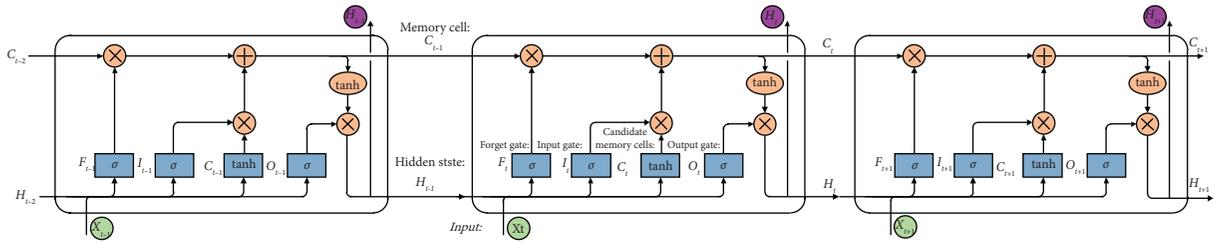


FIGURE 6: Structure of long short-term memory.

**3.2. Model Building and Identification.** In order to verify the effectiveness of situation awareness criteria and LSTM-NN in processing time-series classification tasks, this paper develops a model which contains binary classification and multiclassification, as shown in Figure 7.

**3.2.1. Input Characteristic Data.** In this study, 1197 distracted segments and 2400 ND segments were selected as input data. Each input segment is a 10 s (100 timestamps) time series. In order to reduce the overall error and calculate the statistical characteristics (mean, regression slope, etc.), we take every four steps (four 0.1 s) of the original sequence as a time unit and obtain 25 timestamps. Each timestamp contains 21 indicators.

To eliminate the difference of indexes, the Z-score standardization was utilized to make the indexes conform to normal distribution. Then, 60% of data were randomly selected as the training set, 30% were test set, and 10% were verification set.

**3.2.2. Identification Model Structure.** Referring to [47], the model structure selected in this study is shown in Figure 7. The model consists of two LSTM layers, two full connection

layers, and one dropout layer. As the activation function of the middle layer, ReLU is helpful to solve the problem of network convergence. Sigmoid and softmax are the activation functions of the output layer for binary and multi-recognition models. Adam is adopted as the optimization. The optimal learning rate  $I_r \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$  can be determined by the experiment. The loss function of binary is binary\_crossentropy, and the categorical\_crossentropy is used for multiclassification. The number of neurons in each layer of LSTM is selected from  $\{32, 64, 128, 256\}$ . The superior limit of iterations is set to 1000, and an early stopping mechanism is adopted. After every five epochs, the test results are obtained on the verification set. The training stops when the test error of the verification set increases with the number of epoch times increasing.

**3.2.3. Evaluation Model Performance.** In order to verify the performance of LSTM-NN, this paper uses SVM and AdaBoost as control experiments. There is a category imbalance in data set, so this paper uses F1-score as the evaluation index of classification results. Precision ( $P$ ) is the ratio between the true positives and all the positives. Recall ( $R$ ) is calculated as true positives divided by true positives and false negatives. F1-score is a metric that

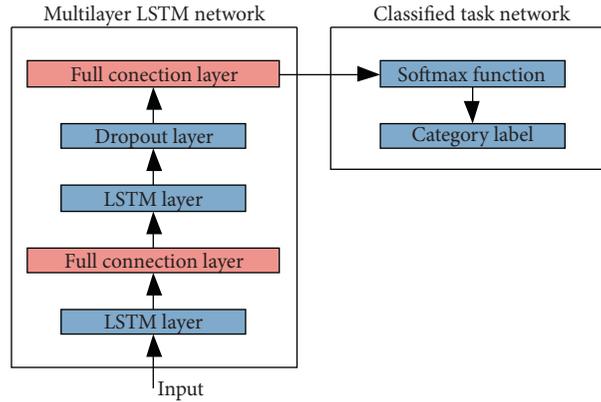


FIGURE 7: Model architecture for time-series classification.

comprehensively considers precision and recall. The calculation formula is as follows:

$$F1 = 2 \times \frac{R \times P}{R + P} \quad (6)$$

#### 4. Results and Discussion

Adopting LSTM-NN to develop a model of binary classification and multiclassification, then utilize SVM and AdaBoost to achieve the same classification task. To avoid overfitting and underfitting, this paper obtains the optimal parameter combination of the three methods, as shown in Table 3. Among Table 3, type 2 (binary classification) is the division of distracted and ND behavior, and type 4 (multiclassification) is the division of driving behavior into which category FDL, FDS, FDU, and ND.

According to the results, ROC curves about binary classification and multiclassification were drawn. The results of the distracted segment were only drawn for the binary classification task. The ROC curve of the multiclassification task included false positive rate, true positive rate, macro average curve, microaverage curve, FDL, FDS, FDU, and ND, as shown in Figure 8. It can be seen intuitively that the ROC curve of each class of LSTM-NN is closer to the (1,1), which is better than SVM and AdaBoost. This difference can be attributed that the parameters of the neural network can handle and learn the high dimensional nonlinear relationship. LSTM-NN takes the information of the front and rear position into account when predicting the state of the next timestamp, which is conducive to the learning of dynamic distracted process.

Table 4 shows the classification results of the model. It can be seen that the F1-score of LSTM-NN for distracted and ND is 89% and 91%, which are better than SVM (77% and 81%) and AdaBoost (79% and 87%). The main reason may be that SVM and AdaBoost lack the ability of dependent long time series, which is one of the reasons why LSTM-NN is used in this paper. When the data of FDL, FDS, FDU, and ND are used, the F1-score of LSTM-NN in identifying three types of distracted situations is 78%, 75%, and 80%. But the accuracy of SVM and AdaBoost in

processing this type of classification task is quite different (the average F1-score is 66% and 71%). It can be seen that LSTM-NN is more capable of recognizing distracted situations than the others. It can also be found that the average F1-score of LSTM-NN in four classifications is reduced by 9% compared to binary classification. On the one hand, the increasing of categories number means the increasing difficulty in correct classification, and more data samples are needed to support model learning. On the other hand, after checking the segments of error classification, it is found that most of the errors come from the false recognition of distracted scenarios. The possible reason is that the data comes from uncontrolled natural driving experiment, and the traffic scene is complex and random. There are some differences in driving behaviors in the same type of distracted scenarios, which leads to a decrease in the accuracy. For example, two segments belonging to FDL have different degrees due to different vehicle speeds and surroundings. The different performance between the above models can be attributed to the better robustness of LSTM-NN and it considers time series in distracted recognition.

Although deep learning has achieved good results, there is still a certain gap between the accuracy of the model in this paper and the related studies. There may be several reasons for this.

Firstly, the data source is different from the experimental environment. This paper uses the noninvasive features of natural driving data. But most of the related research studies are carried out in the driving simulator and controlled experiment, which may reduce the randomness of distraction and weaken the distracted effect on vehicle kinematics characteristics.

The second is the recognition indexes that are different. This paper only selects indicators from the vehicle movement. The related researches include driver's posture such as head angle. If such index fluctuates violently, the driver is regarded as distracted. However, it is worthy to point out that there are a lot of sightline deviations or head rotations in the actual driving process, which belongs to a normal phenomenon in the driving process instead of inattention. So this kind of distraction situation has been deleted by video review in the study.

TABLE 3: Related parameters of classification models.

LSTM-NN									
Type	Neurons numbers	Iterations	Batch size	Output layer activation function	Optimizer	Learning rate	Loss function		
2	64	100	32	Sigmoid	Adam	$10^{-3}$	binary_crossentropy		
4	64	200	32	Softmax		$10^{-2}$	categorical_crossentropy		

SVM		
Type	Main parameters	Optimum parameters
2	C, gamma, and kernel	$C = 10^3$ , $\gamma = 10^{-3}$ , and kernel = rbf
4		$C = 10^2$ , $\gamma = 10^{-2}$ , and kernel = rbf

AdaBoost		
Type	Main parameters	Optimum parameters
2	Learning rate and n_estimation	Learning rate = 0.1 and n_estimation = 50
4		Learning rate = 0.15 and n_estimation = 100

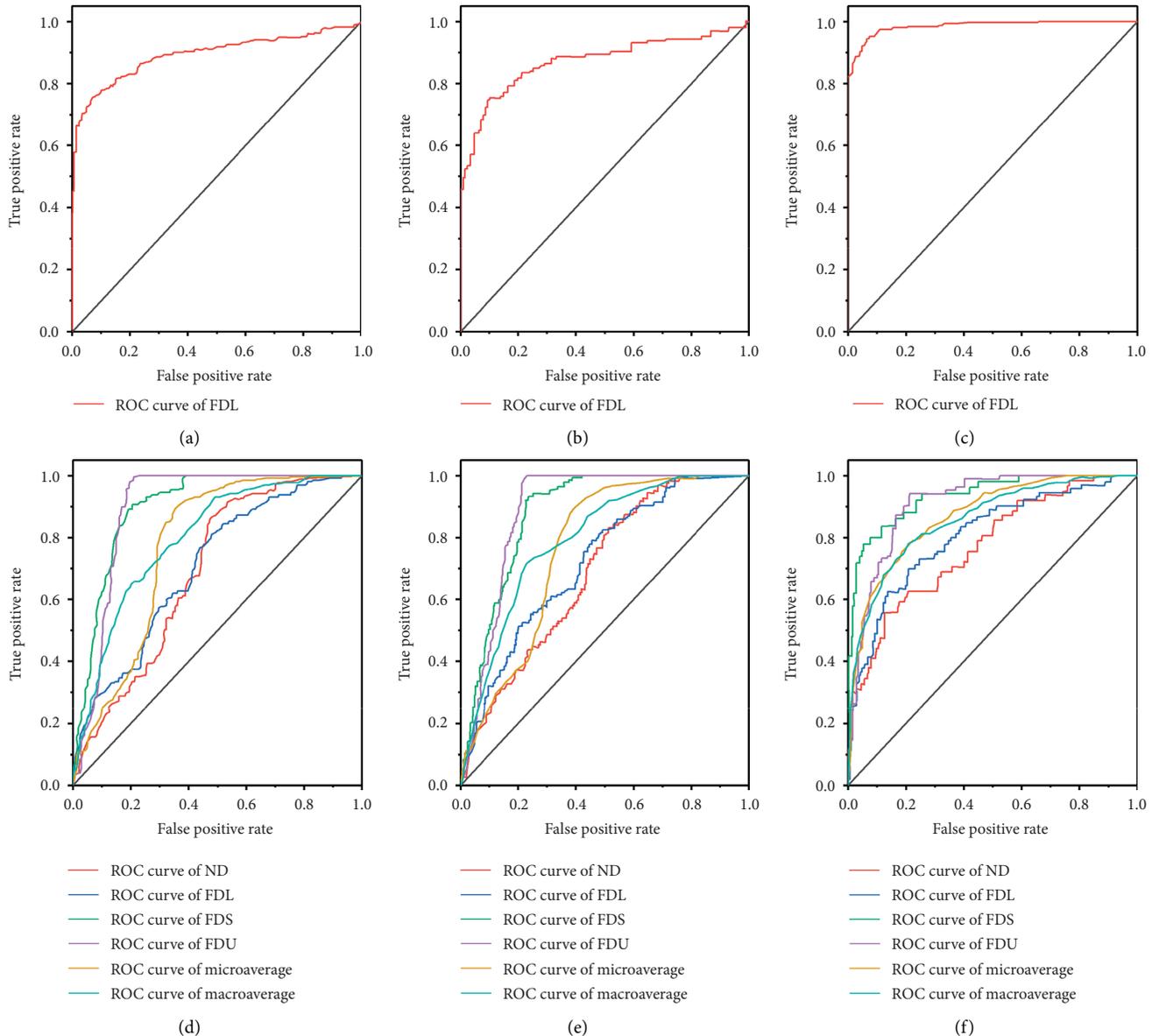


FIGURE 8: Classification results of the models. (a) AdaBoost of binary categories, (b) SVM of binary categories, (c) LSTM-NN of binary categories, (d) AdaBoost of multicategories, (e) SVM of multicategories, and (f) LSTM-NN of multicategories.

TABLE 4: Results of distraction classification.

LSTM-NN																
Type	Accuracy				Recall				FI-score				G-mean			
2	Distraction		Normal													
	88%		92%		90%		89%		89%		91%		89%		90%	
4	FDL	FDS	FDU	ND												
	74%	71%	77%	91	82%	79%	84%	92	78%	75%	80%	91	77%	76%	80%	91%
SVM																
Type	Accuracy				Recall				FI-score				G-mean			
2	Distraction		Normal													
	76%		82%		78%		80%		77%		81%		76%		81%	
4	FDL	FDS	FDU	ND												
	72%	69%	64%	80	56%	66%	68%	85	64%	67%	66%	83	63%	67%	66%	82%
AdaBoost																
Type	Accuracy				Recall				FI-score				G-mean			
2	Distraction		Normal													
	77%		87%		81%		86%		79%		87%		79%		86%	
4	FDL	FDS	FDU	ND												
	73%	71%	68%	86	69%	75%	70%	85	71%	73%	69%	86	72%	72%	68%	86%

## 5. Conclusion

Based on the natural driving data of real vehicles from SHRP2 NDS, this paper focuses on the relationship between vehicle kinematic characteristics and distracted driving state and proposes a recognition method of distraction. The main research conclusions are as follows.

This paper attempts to extract distracted events in uncontrolled natural driving environment. Compared with related researches through a driving simulator or arranged experiments, an extraction method based on situation awareness is proposed. This method provides a new data preparation effectively for the study of distraction behavior.

By using the noninvasive vehicle kinematic features, GBDT-RFE and RF-RFE are used to select the indexes with high importance. The way not only reduces the interference of redundant indexes but also verifies the effectiveness of the motion features as the distracted recognized conditions.

In this paper, LSTM-NN deep learning is used to classify distracted states. Results show that the model is better than other models in dealing with binary classification and multiclassification. It is proved that distraction is a dynamic fluctuation process with time, and LSTM-NN can learn the information of time series before and after and detect the distracted behavior reliably.

This paper adopts noninvasive vehicle kinematics feature data to detect off-line driving distraction and proposes a method based on situation awareness. The distracted data set is obtained by inferring whether the driver is distracted by the vehicle motion situation. Then the distracted time sequence segments are identified by using the deep learning model. In the next step, driver's attention information is extracted by image recognition technology combined with in-vehicle video data. Real-time online recognition of distracted state based on multisource unstructured data will be further studied. This part of the work is in progress.

## Data Availability

All data included in this study are available upon request by contacting the corresponding author.

## Conflicts of Interest

The authors declare that they have no conflicts of interest in this work.

## Authors' Contributions

Zhi-Qiang Liu performed conceptualization and supervision. Shi-Heng Ren contributed to investigation, formal analysis, and writing original draft. Man-Cai Peng carried out visualization, formal analysis, writing original draft, and methodology.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (61403172). The authors thank all the participants, school administrators, and local government who helped facilitate our intervention work.

## References

- [1] NHTSA, *Distracted Driving 2015*, NHTSA, NY, USA, Dot Hs 812 132 DOT HS 811 379, 2017.
- [2] H. Singh and A. Kathuria, "Analyzing driver behavior under naturalistic driving conditions: a review," *Accident Analysis & Prevention*, vol. 150, p. 105908, 2021.
- [3] K. J. Parnell, J. Rand, and K. L. Plant, "A diary study of distracted driving behaviours," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 74, pp. 1–14, 2020.
- [4] M. A. Regan and C. Hallett, "Driver distraction: definition, mechanisms, effects, and mitigation," in *Handbook of Traffic Psychology*, Elsevier, Amsterdam, Netherlands, 2011.

- [5] Y. Liang and J. D. Lee, "Combining cognitive and visual distraction: less than the sum of its parts," *Accident Analysis & Prevention*, vol. 42, no. 3, pp. 881–890, 2010.
- [6] J. Gao and G. A. Davis, "Using naturalistic driving study data to investigate the impact of driver distraction on driver's brake reaction time in freeway rear-end events in car-following situation," *Journal of Safety Research*, vol. 63, pp. 195–204, 2017.
- [7] S. W. Savage, D. D. Potter, and B. W. Tatler, "The effects of cognitive distraction on behavioural, oculomotor and electrophysiological metrics during a driving hazard perception task," *Accident Analysis & Prevention*, vol. 138, p. 105469, 2020.
- [8] P. D'Addario and B. Donmez, "The effect of cognitive distraction on perception-response time to unexpected abrupt and gradually onset roadway hazards," *Accident Analysis & Prevention*, vol. 127, pp. 177–185, 2019.
- [9] A. S. Le, T. Suzuki, and H. Aoki, "Evaluating driver cognitive distraction by eye tracking: from simulator to driving," *Transportation Research Interdisciplinary Perspectives*, vol. 4, p. 100087, 2020.
- [10] D. B. Kaber, Y. Liang, Y. Zhang, M. L. Rogers, and S. Gangakhedkar, "Driver performance effects of simultaneous visual and cognitive distraction and adaptation behavior," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 15, no. 5, pp. 491–501, 2012.
- [11] L. Li, B. Zhong, C. Hutmacher, Y. Liang, W. J. Horrey, and X. Xu, "Detection of driver manual distraction via image-based hand and ear recognition," *Accident Analysis & Prevention*, vol. 137, p. 105432, 2020.
- [12] R. Broström, P. Bengtsson, and M. L. Aust, "Individual glance strategies and their effect on the NHTSA visual manual distraction test," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 36, pp. 83–91, 2016.
- [13] W. Qu, Y. Ge, Y. Guo, X. Sun, and K. Zhang, "The influence of WeChat use on driving behavior in China: a study based on the theory of planned behavior," *Accident Analysis & Prevention*, vol. 144, p. 105641, 2020.
- [14] Z. Li, S. Bao, I. V. Kolmanovsky, and X. Yin, "Visual-Manual distraction detection using driving performance indicators with naturalistic driving data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2528–2535, 2018.
- [15] E. Muhrer and M. Vollrath, "The effect of visual and cognitive distraction on driver's anticipation in a simulated car following scenario," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 14, no. 6, pp. 555–566, 2011.
- [16] H. Grahn and T. Kujala, "Impacts of touch screen size, user interface design, and subtask boundaries on in-car task's visual demand and driver distraction," *International Journal of Human-Computer Studies*, vol. 142, p. 102467, 2020.
- [17] Y. Liao, S. E. Li, G. Li, W. Wang, B. Cheng, and F. Chen, "Detection of driver cognitive distraction: an SVM based real-time algorithm and its comparison study in typical driving scenarios," in *Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV)*, Gothenburg, Sweden, June 2016.
- [18] T. Chihara, F. Kobayashi, and J. Sakamoto, "Evaluation of mental workload during automobile driving using one-class support vector machine with eye movement data," *Applied Ergonomics*, vol. 89, p. 103201, 2020.
- [19] T. Liu, Y. Yang, G.-B. Huang et al., "Cluster regularized extreme learning machine for detecting mixed-type distraction in driving," in *Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, September 2015.
- [20] U. Fugiglando, E. Massaro, P. Santi et al., "Driving behavior analysis through CAN bus data in an uncontrolled environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 737–748, 2019.
- [21] K. Iio, X. Guo, and D. Lord, "Examining driver distraction in the context of driving speed: an observational study using disruptive technology and naturalistic data," *Accident Analysis & Prevention*, vol. 153, p. 105983, 2021.
- [22] R. Arvin and A. J. Khattak, "Driving impairments and duration of distractions: assessing crash risk by harnessing microscopic naturalistic driving data," *Accident Analysis & Prevention*, vol. 146, p. 105733, 2020.
- [23] M. Risteska, D. Kanaan, B. Donmez, and H.-Y. Winnie Chen, "The effect of driving demands on distraction engagement and glance behaviors: results from naturalistic data," *Safety Science*, vol. 136, p. 105123, 2021.
- [24] X. Wang, R. Xu, A. Asmelash, Y. Xing, and C. Lee, "Characteristics of driver cell phone use and their influence on driving performance: a naturalistic driving study," *Accident Analysis & Prevention*, vol. 148, p. 105845, 2020.
- [25] D. Vetturi, M. Tiboni, G. Maternini, and M. Bonera, "Use of eye tracking device to evaluate the driver's behaviour and the infrastructures quality in relation to road safety," *Transportation Research Procedia*, vol. 45, pp. 587–595, 2020.
- [26] C. D. D. Cabrall, R. Happee, and J. C. F. de Winter, "Prediction of effort and eye movement measures from driving scene components," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 68, pp. 187–197, 2020.
- [27] C.-T. Lin, S.-A. Chen, L.-W. Ko, and Y.-K. Wang, "EEG-based brain dynamics of driving distraction," in *Proceedings of the 2011 International Joint Conference on Neural Networks*, San Jose, CA, USA, August 2011.
- [28] E. Ohn-Bar, S. Martin, A. Tawari, and M. M. Trivedi, "Head, eye, and hand patterns for driver activity recognition," in *Proceedings of the 2014 22nd International Conference on Pattern Recognition*, Stockholm, Sweden, August 2014.
- [29] D. Anitta and A. Annis Fathima, "Human head pose estimation based on HF method," *Microprocess. Microsyst.* vol. 82, 2021.
- [30] M. Van Ly, S. Martin, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors," in *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV)*, Gold Coast, QLD, Australia, June 2013.
- [31] Y. Ma, K. Tang, S. Chen, A. J. Khattak, and Y. Pan, "On-line aggressive driving identification based on in-vehicle kinematic parameters under naturalistic driving conditions," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 554–571, 2020.
- [32] N. Li and C. Busso, "Using perceptual evaluation to quantify cognitive and visual driver distractions," in *Smart Mobile In-Vehicle Systems: Next Generation Advancements*, pp. 183–207, Springer, New York, NY, USA, 2014.
- [33] G. K. Kountouriotis, P. Spyridakos, O. M. J. Carsten, and N. Merat, "Identifying cognitive distraction using steering wheel reversal rates," *Accident Analysis & Prevention*, vol. 96, pp. 39–45, 2016.
- [34] M. Miyaji, H. Kawanaka, and K. Oguri, "Driver's cognitive distraction detection using physiological features by the adaboost," in *Proceedings of the 2009 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, USA, October 2009.

- [35] A. Azman, Q. Qinggang Meng, and E. Edirisinghe, "Non intrusive physiological measurement for driver cognitive distraction detection: eye and mouth movements," in *Proceedings of the 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, Chengdu, China, August 2010.
- [36] D. W. Ryu, H. B. Jeong, S. H. Lee, W.-S. Lee, and J. H. Yang, "Development of driver-state estimation algorithm based on Hybrid Bayesian Network," in *Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV)*, Seoul, Korea (South), June 2015.
- [37] Y. Liang and J. D. Lee, "A hybrid Bayesian Network approach to detect driver cognitive distraction," *Transportation Research Part C: Emerging Technologies*, vol. 38, pp. 146–155, 2014.
- [38] M. Wollmer, C. Blaschke, T. Schindl et al., "Online driver distraction detection using long short-term memory," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 574–582, 2011.
- [39] C. N. Pope, T. R. Bell, and D. Stavrinou, "Mechanisms behind distracted driving behavior: the role of age and executive function in the engagement of distracted driving," *Accident Analysis & Prevention*, vol. 98, pp. 123–129, 2017.
- [40] O. Nakayama, T. Futami, T. Nakamura, and E. R. Boer, "Development of a steering entropy method for evaluating driver workload," *SAE Technical Paper Series*, vol. 108, pp. 1686–1695, 1999.
- [41] R. Van Der Horst and J. Hogema, "Time-to-collision and collision avoidance systems," in *Proceedings of the 6th ICTCT workshop*, Salzburg, Austria, 1993.
- [42] S. J. Kass, K. S. Cole, and C. J. Stanny, "Effects of distraction and experience on situation awareness and simulated driving," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 10, no. 4, pp. 321–329, 2007.
- [43] D. Cornacchio, A. L. Sanchez, T. Chou, and J. S. Comer, "Cognitive-behavioral therapy for children and adolescents," in *The Science of Cognitive Behavioral Therapy*, Elsevier, Amsterdam, Netherlands, 2017.
- [44] G. Panagopoulos and I. Pavlidis, "Forecasting markers of habitual driving behaviors associated with crash risk," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 2, pp. 841–851, 2020.
- [45] Z. Xu, G. Huang, K. Q. Weinberger, and A. X. Zheng, "Gradient boosted feature selection," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, August 2014.
- [46] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [47] X. Huang, J. Sun, and J. Sun, "A car-following model considering asymmetric driving behavior based on long short-term memory neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 346–362, 2018.