IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 6, 2019, accepted December 12, 2019, date of publication December 16, 2019, date of current version December 27, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2960157*

# Drowsy Driving Detection Based on Fused Data and Information Granulation

**YAN WANG[1,2], LISHENG JIN[1], KEYONG LI[3], BAICANG GUO[1], YI ZHENG[1], AND JIAN SHI[4]**

[1]Transportation College, Jilin University, Changchun 130022, China
[2]School of Science, Jilin Institute of Chemical Technology, Jilin 132022, China
[3]Shanghai Huawei Technologies Company, Ltd., Shanghai 201206, China
[4]College of Communication Engineering, Jilin University, Changchun 130022, China

Corresponding author: Lisheng Jin (jinls@jlu.edu.cn)

This work was supported by the National Natural Science Foundation under Grant 51575229.

**ABSTRACT** To detect drowsy driving accurately, this paper collects the characteristic parameters of driver's operating behavior and vehicle's running state through simulation experiments. Then, the factor analysis was adopted to reduce the dimensionality of the characteristic parameters, and the composite factor scores was computed under both normal and drowsy states, forming a time series. Next, the time series of composite factor scores was divided into information granules and the particle swarm optimization (PSO) was implemented to optimize the dynamic time window for the characteristic parameters. After that, the mean and standard deviation were computed for each composite factor score in each sub-time window, and fused into one data. Finally, a drowsy driving model was established with LIBSVM based on the fused data. The experimental results demonstrate that our model achieved an accuracy of 86.47% in detecting drowsy driving. The research finding shed new lights on the detection of drowsy driving.

**INDEX TERMS** Drowsy driving, information granulation, support vector machine (SVM), dynamic time window.

## I. INTRODUCTION

The driver, a key player in the road traffic system, has long been considered as a primary factor in traffic accidents. During driving, potential conflicts may arise between the driver and other drivers and pedestrians. The driver must notice the signs of conflict, and take quick and proper measures to avoid the conflict, achieving safe and smooth driving. This requires the driver to be energetic and stay focused, rather than be sidetracked. In the case of drowsy driving, however, the driver will face a decline in physical strength and find it hard to keep attention, make a correct judgment or execute proper operations. The common causes of drowsy driving include lack of adequate sleep, extended driving time, medications, or drinking alcohol. The driving fatigue is directly related to sleep quality. Sleep is the best way for drivers to ease fatigue and restore energy.

Driving is more of mental work than physical work. When he/she is driving, the driver must keep an eye on the changing road and traffic conditions, while predicting the changes in a certain range ahead of the vehicle. After driving for a long distance or long time, the concentrated efforts will lead to muscle tension, poor blood circulation, and insufficient oxygen supply. If the driver takes sleeping pills or analgesics, the medicines will act on the central nervous system and cause drowsiness. If the driver drinks alcohol-containing beverages, drowsiness will also occur under the effect of alcohol.

A drowsy driver cannot accurately perceive or handle traffic conditions in time. Therefore, many major traffic accidents are potentially caused by drowsy driving. This dangerous driving mode is 4~6 times more likely to induce traffic accidents than normal driving [1]. In fact, drowsy driving contributes to about 20% of all traffic accidents and more than 40% of major traffic accidents each year [2], [3]. According to the estimation by the National Highway Traffic Safety Administration (NHTSA) of the US, the police reported more than 72,000 crashes involves drowsy drivers in 2015 alone, leading to 41,000 injuries and more than 800 deaths [4].

The danger of drowsy driving calls for attention from both the driver and the authority of traffic management. However, the efforts from the two parties are hindered by several factors. First, a drowsy driver is slow in response and often underestimates his/her fatigue level. Second, many drivers, driven by economic interests, often continue to drive

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

for a long time after feeling drowsy, ignoring the traffic laws. Third, drowsy driving is very concealed and difficult to be monitored by the authority. Thus, technical measures should be adopted to control drowsy driving, in addition to policy making. Most importantly, new techniques must be developed to accurately identify the state of drowsy driving, alerting both driver and the authority to take action. Therefore, it is of great significance to detect the state of the driver in real time and issue an effective warning of drowsy driving.

Drowsy driving detection can be divided into two categories: subjective evaluation method and objective detection method.

If performed by the driver, the subjective evaluation usually uses drowsiness scales like Stanford Sleepiness Scale and the Piper Fatigue Scale; the drowsiness levels are basically set up through questionnaire surveys [5]. If performed by others, the subjective evaluation is often realized through expert scoring of videos on the driver's facial expressions. By this method, each video is divided into several segments, and several experts are invited to identify the drowsiness level of the driver according to the facial expressions, such as yawning, eye shift and head posture [6].

The objective detection means finding out the characteristic parameters of drowsy driving according to the physiology, psychology and operating behavior of the driver. The characteristic parameters can be extracted either from bioelectrical signals like electrocardiogram (EEG), electrocardiogram (ECG), electroencephalogram (EOG) and respiratory signals, or from facial features like eye movements, face expression and head position.

The early studies that detect drowsy driving based on bioelectrical signals have mostly relied on the EEG signal. Yeo *et al.* [7] established a support vector machine (SVM) model based on drivers' EEG signals, and accurately recognized 95% of drowsing driving states. Lin *et al.* [8], [9] designed a real-time wireless EEG-based brain-computer interface system and an EEG-based self-organizing fuzzy neural system to predict drowsy driving. Later, many scholars combined the EEG, ECG, and EOG to extract the characteristic parameters. For instance, Fujiwara *et al.* [10] proposed a driver drowsiness detection algorithm based on heart rate variability (HRV) analysis using the ECG and validated the algorithm in comparison to EEG-based sleep scoring. Wang *et al.* [11] performed a fusion entropy analysis to detect driver drowsiness. Ahn *et al.* [12] collected multimodal EEG/ECG/EOG and functional near-infrared spectroscopy (fNIRS) data [12], and then developed algorithms to identify neurophysiological correlations of drivers' mental states. Guede-Fernández *et al.* [13] designed a drowsiness detection method based on changes in respiratory signals, which processes the respiratory signals in real time to classify the driver into drowsy state or awake state. The existing bioelectrical signal-based drowsiness detection methods need to measure the signals in contact mode, which has many limitations in actual applications.

The existing studies that detect drowsy driving based on facial features have utilized various features, ranging from eye blinking rate [14], slow eye movements [15], the percentage of eye closure [16]–[19] to face expressions. Saradadevi [20] and Abtahi *et al.* [21] tracked the mouth of drivers and recognized yawning, thus detecting driver drowsiness. Jie *et al.* [22] presented an automatic yawning detection approach that extracts the geometric and appearance features of both mouth and eye regions, which successfully detected 95% of yawns, whether covered or not covered by hands. In general, facial feature-based detection methods for drowsy driving need to track and process facial images, posing a challenge to real-time performance.

In fact, the drowsiness of drivers can be identified based on unnatural operating behavior like the abnormal angle and angular speed of the steering wheel. McDonald *et al.* [23] put forward a method to detect drowsiness based on lane departures. The method uses unfiltered data of steering wheel angle and a random forest algorithm. Li *et al.* [24] developed an online drowsiness detection system to monitor driver fatigue level under real driving conditions, using the steering wheel angle collected by sensors mounted on the steering lever. Arefnezhad *et al.* [25], Chai *et al.* [26] and Schwarz *et al.* [27] also extracted features from the data on the steering wheel, and treated them as important indices for drowsy driving detection.

Owing to the diversity of data, it is of great significance to fuse different types of information. Xiao *et al.* [28], [29] proposed a hybrid approach to describe the trend of information dissemination in social networks, which couples multidimensional user attributes, evolutionary games, and the traditional susceptible-infected-recovered epidemic model. Jin *et al.* [30] and Guo *et al.* [31] fused the data on the eye movements of drivers and those on vehicle running status, the index system was established to evaluate secondary task driving based on the fused data, and analyzed the importance of the system to driving safety. To improve the accuracy of drowsy driving detection, many other methods have been developed based on the fusion of yawn and head movement [32], or the fusion of multiple types of features [33]–[36], e.g. the facial features of the driver and the steering features of the vehicle.

In many studies, the index data used to detect drowsy driving are fused from various types of features, each of which has its inherent features. In this case, it is a key issue to select a suitable size of the time window for operation. However, little attention has been paid to how time window size affects the detection of drowsy driving. In most researches, the time window of an index is usually selected empirically. For example, Eskandarian *et al.* [37] directly used a 15s-long time window for steering wheel data. Zhang *et al.* [38] adopted a uniform 30s-long time window for different indices used to detect drowsing driving, namely, eye movement, steering wheel operation, and vehicle's running state. Some scholars have selected time windows of different sizes, rather than a fixed time window. For instance, Niu [39] analyzed the angle of the steering wheel with time windows of four sizes, i.e. 5s, 10s, 15s, and 20s. Wang *et al.* [40] explored the effects of different combinations of input parameters with different sizes of time windows (4s, 10s, 20s, 30s and 60s) on driving detection based on the random forest; the parameters include lateral acceleration, longitudinal acceleration, and steering angle. To sum up, the same index has been studied with a fixed time window or time windows of different sizes, while different indices have been discussed with a fixed time window. Compared with a dynamic time window, the fixed time window is not conducive to mining the intrinsic features of each index.

In drowsy driving detection, the extracted indices form a special time series, which can be processed by information granulation put forward by Zadeh in 1979 [41]. Information granulation splits a piece of information into several granules and processes them separately with a dynamic time window. Later, Pedrycz developed the information granulation of time series, which becomes increasingly popular in recent years [42]–[46].

Factor analysis [47], proposed by Spearman in 1904, is a dimensionality reduction method that extracts common factors from a number of variables. In factor analysis, the original variables that are closely correlated with each other are allocated to the same group, such that the intra-group correlation is greater than inter-group correlation. Each group represents a basic structure and has a hidden parameter called the common factor. The common factor helps to reduce the dimensionality of the collected indices and obtain a composite factor score of each index.

Drawing on the relevant literature, this paper carries out experiments on experienced young and middle-aged drivers on a driving simulation platform, and collects characteristic parameters of driver's operating behavior (e.g. steering wheel angle, steering wheel angular speed, throttle pedal opening, and brake pedal opening) and those of the vehicle's running state (e.g. lateral speed, longitudinal speed, lateral acceleration, longitudinal acceleration, yaw angle, yaw angular speed, and yaw angular acceleration). The collected parameters were fused and used to establish a drowsy driving model was established with LIBSVM [48]. There are three main contributions of this research:

(1) Since the accuracy of drowsy driving detection hinges on the indicative parameters, whose intrinsic features are not easy to explore, the authors proposed a novel method based on the fused data and information granulation. This method guarantees the comprehensive performance of the parameters and fully considers their respective features through data fusion and information granulation.

(2) Considering the sheer number of indicative parameters of drowsy driving, the factor analysis was introduced to reduce the dimensionality, and output composite factor scores of the parameters. The composite factor scores form a time series. Through the analysis, the main influencing factors and their impacts were identified, and the score of common factors was obtained, laying the basis for subsequent data fusion.

(3) A fixed time window cannot track the changes in parameter data timely, making it impossible to interpret the data rationally. To solve the problem, the information granulation of time series was adopted to split the time series of common factors into information granules. Then, the particle swarm optimization (PSO) [49] was performed to optimize the dynamic time window. Compared with fixed time windows, the dynamic time window brought high data fitness and significant indices in each sub-time window. This is a novel method for selecting the size of time windows for indicative parameters in drowsy driving detection.

The remainder of this paper is organized as follows: Section II introduces the methods adopted for this research and explains the model establishment; Section III presents the experiments on our model and analyzes the experimental results; Section IV further discusses our model; Section 5 sums up the main findings of this research.

## II. METHODOLOGY

This section mainly introduces the establishment of the drowsy driving model based on methods like factor analysis, information granulation of time series and the SVM.

### A. FACTOR ANALYSIS

Proposed by Spearman in 1904, factor analysis [47] is a dimensionality reduction method that extracts common factors from several variables. In factor analysis, the original variables that are closely correlated with each other are allocated to the same group, such that the intra-group correlation is greater than inter-group correlation. Each group represents a basic structure and has a hidden parameter called the common factor. The most popular model for factor analysis is as follows:

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \varepsilon_1 \\ X_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + \varepsilon_2 \\ \vdots \\ X_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + \varepsilon_p \end{cases} \tag{1}$$

Equation (1) can be simplified as:

$$X_{p\times 1} = A_{p\times m}F_{m\times 1} + \varepsilon_{p\times 1} \tag{2}$$

where, $X = (X_1, \cdots, X_p)^T$ is a $p$-dimensional random vector consisting of $p$ measurable indices; $F = (F_1, \cdots, F_m)^T$ are the common factors of $X$; $a_{ij}$ is the load of variable $i$ on common factor $j$; $A$ is the factor load matrix; $\varepsilon = \varepsilon_1, \cdots, \varepsilon_p$ are the special factors of $X$.

The factor analysis model is subjected to the following constraints:

The number of common factors must be less than or equal to the number of measurable indices:

$$m \leq p$$

The common factor $F$ is not correlated with the special factor $\varepsilon$:

$$Cov(F, \varepsilon) = 0$$

The common factors $F_1, \cdots, F_m$ are not correlated with each other, and their variances are the same (one):

$$D(F) = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} = I_m \tag{3}$$

The special factors $\varepsilon_1, \cdots, \varepsilon_p$ are not correlated with each other, and their variances are different:

$$D(\varepsilon) = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_p^2 \end{bmatrix} \tag{4}$$

### B. INFORMATION GRANULATION

There are three theories on information granulation: fuzzy set theory [41], rough set theory [50] and quotient space theory [51].

The information granulation of time series is mainly implemented in the following steps:

*Step 1:* Creating information granules

For the n-element time series $X = \{x_i | i = 1, 2, \cdots, n\}$, initialize $q$ time windows $T_1, T_2, \cdots, T_q$ to divide the time series into $q$ subseries $D_1, D_2, \cdots, D_q$ according to the rational granulation criterion [42], creating $q$ information granules $\Omega_1, \Omega_2, \cdots, \Omega_q$.

*Step 2:* Identifying information granules

Find the optimal information granule $\Omega_i(\alpha)$, $\alpha \in [0, 1]$ by computing the optimal lower and upper bounds $a_i$ and $b_i$ of information granule $\Omega_i$ in time window $T_i$:

$$\begin{aligned} a_i &= \arg\max_{a \le med(\Omega_i)} \{card\, \{x_k \in \Omega_i, a \le x_k \le med(\Omega_i)\} \\ &\quad \times \exp(-\alpha\, |med(\Omega_i) - a|)\} \end{aligned} \tag{5}$$

$$\begin{aligned} b_i &= \arg\max_{b \ge med(\Omega_i)} \{card\, \{x_k \in \Omega_i, med(\Omega_i) \le x_k \le b\} \\ &\quad \times \exp(-\alpha\, |b - med(\Omega_i)|)\} \end{aligned} \tag{6}$$

where, card $(x_k \in \Omega_i)$ and med $(\Omega_i)$ are the number and median of the data in information granule $\Omega_i$; $\alpha \in [0, 1]$ is the given parameter.

*Step 3:* Computing information granule indices

Calculate the index $Vol(\Omega_i)$ of the information granule $\Omega_i$ in time window $T_i$:

$$Vol(\Omega_i) = |T_i| \int_{x_{min}}^{x_{max}} |\Omega_i(z)| dz \tag{7}$$

where, $|\cdot|$ is the length or size of the corresponding quantity; $x_{\max}$ and $x_{\min}$ are the maximum and minimum amplitudes of the subseries in the time window.

*Step 4:* Optimizing time window size

According to the rational granulation criterion, the information granule index $Vol(\Omega_i)$ is negatively correlated with the number of data in the information granule $\Omega_i$. Then, the optimal length of the time window $T_1^*, T_2^*, \cdots, T_q^*$ can be obtained by:

$$\min_{T_1, T_2, \cdots, T_q} \sum\nolimits_{i=1}^{q} Vol(\Omega_i) \tag{8}$$

### C. THE SVM

The SVM is a supervised machine learning method proposed by Vapnik based on the Vapnik–Chervonenkis (VC) dimension and structural risk minimization (SRM) [52]. The core idea of the SVM is to find the optimal classification hyperplane in the sample space, and determine the maximum interval to divide the samples into two categories. The specific process of the binary classification depends on the linearity of the samples.

#### 1) BINARY CLASSIFICATION OF LINEAR SAMPLES

In a d-dimensional space, the binary classification of sample set $(x_i, y_i)$ $(x \in R^d, y \in \{+1, -1\}, i = 1, 2, \ldots, n)$ can be defined as:

$$g(x) = w^T \cdot x + b \tag{9}$$

The classification hyperplane can be obtained by:

$$w^T \cdot x + b = 0 \tag{10}$$

In real-world scenarios, the samples are often inseparable linearly. To solve the problem, a penalty factor $C \sum_{i=1}^{n} \xi_i$ for linear separability was introduced to convert the optimization of the classification plane into:

$$\begin{aligned} &\min \left( \frac{1}{2} ||w||^2 + C \sum_{i=1}^{n} \xi_i \right) \\ &s.t. \begin{cases} y_i(w^T \cdot x + b) \ge 1 - \xi, & i = 1, 2, \ldots, n, \\ \xi_i \ge 0, & i = 1, 2, \ldots, n. \end{cases} \end{aligned} \tag{11}$$

Then, the Lagrange function based on (11) can be solved by:

$$\begin{aligned} L(w, b, \alpha) &= \frac{1}{2}(w^T \cdot w) + C \sum_{i=1}^{n} \xi_i \\ &\quad - \sum_{i=1}^{n} \alpha_i [y_i(w^T \cdot x_i + b) - 1 + \xi_i] \end{aligned} \tag{12}$$

where, $\alpha_i \ge 0$ is the Lagrange coefficient. The following can be derived from the partial differentials w and b in (12):

$$w = \sum_{i=1}^{n} \alpha_i x_i y_i \tag{13}$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \tag{14}$$

Substituting (13) and (14) into (10), the original problem can be transformed into:

$$\begin{aligned} &\max_{\alpha} L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ &s.t. \begin{cases} 0 \le \alpha_i \le C, & i = 1, 2, \ldots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases} \end{aligned} \tag{15}$$

Then, the function of the optimal classification hyperplane can be obtained by solving (15):

$$f(x) = \text{sgn}\{\sum_{i=1}^{n} \alpha_i y_i K(x_i \cdot x_j) + b\} \tag{16}$$

2) BINARY CLASSIFICATION OF NONLINEAR SAMPLES

The SVM relies on the kernel function to solve the binary classification of nonlinear samples:

$$K(x_i \cdot x_j) = \phi(x_i)\phi(x_j) \tag{17}$$

According to (17), the binary classification of nonlinear samples can be converted into:

$$\max_{\alpha} L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j)$$
$$s.t. \begin{cases} 0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, n, \\ \sum_{i=1}^{n} \alpha_i y_i = 0. \end{cases} \tag{18}$$

### *D. THE PROPOSED METHOD*

Fig. 1 shows the structural of the proposed method for drowsy driving detection. Obviously, the proposed method mainly consists of three stages.

In Stage 1, the characteristic parameters of driver's operating behavior and those of the vehicle's running state were collected on the driving simulation platform. The collected parameters were taken as the indicative parameters to be preprocessed to form the drowsy driving dataset. In the experiment, every driver is required to drive the test vehicle twice, once in normal state and the other in drowsy state.

In Stage 2, the indicative parameters were preprocessed and fused. Firstly, the factor analysis was performed to reduce the dimensions of the indicative parameters in normal and drowsy states, and to determine the composite factor scores of them, respectively. Secondly, the information granulation of time series was adopted to split the series of composite factor scores into information granules, and the PSO was performed to optimize the dynamic time window. Finally, the mean and standard deviation of composite factor scores were computed and fused into one data in the optimal dynamic time window.

In Stage 3, a drowsy driving detection model was established based on LIBSVM. The fused data were randomly divided into a training set and a test set. Then, the model was trained to optimize the penalty coefficient and kernel function parameter with the training set. Finally, the established model was tested with the test set.

The pseudo code of the drowsy driving detection model is shown in Algorithm 1. The time complexity of the model comes from three parts: dimensionality reduction, the PSO of dynamic time window, and the model training and test. Thus, the time complexity of Algorithm 1 can be expressed as $T = T_{dr} + T_{dtw} + T_{dm} \sim O(n^2)$, where $T_{dr} = O(n)$, $T_{dtw} = O(n^2)$ and $T_{dm} = O(n^2)$ are the time complexities of the three parts, respectively. In fact, the problem scale was decreased through dimensionality reduction, and further reduced from n to q through the optimization of dynamic time window. In general, Algorithm 1 has an acceptable level of time complexity.

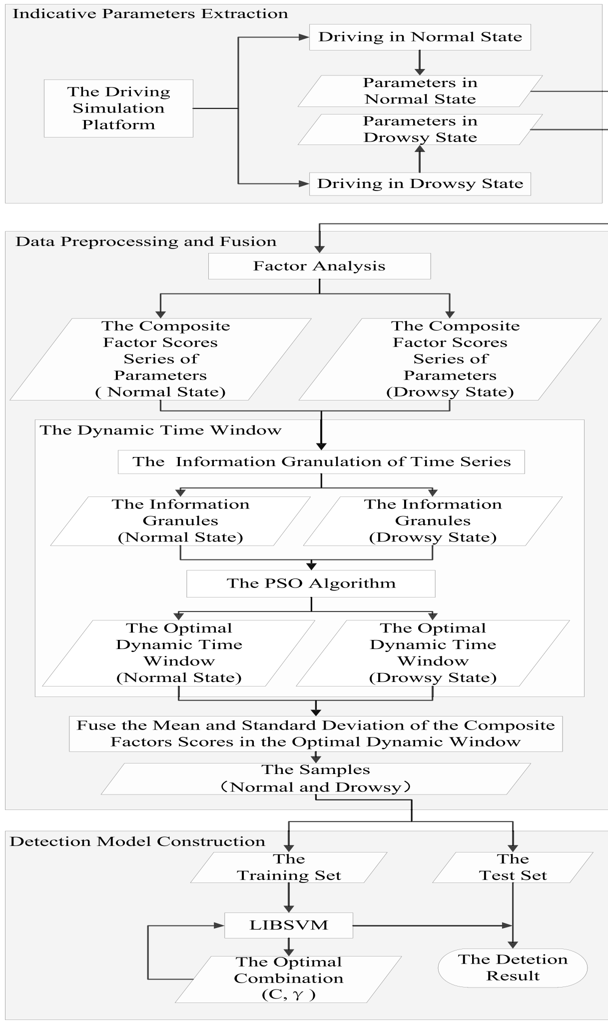


**FIGURE 1. The framework of the proposed method.**

## III. EXPERIMENTS AND RESULTS

### *A. EXPERIMENTS*

1) SUBJECTS

Female drivers generally have weaker motor perception skills, and thus poorer control of the vehicle, than their male counterparts [53]. Therefore, female drivers are more likely to face abnormal changes in the running state of the vehicle. Our drowsy driving detection is based on driver operation behavior and vehicle running state. To minimize the influence of personal factors, about 25% of the subjects of our experiments are female drivers.

In total, 16 experienced drivers, aged between 24 and 4, including 4 females and 12 males, were selected for our experiments. The mean age was 29.7 years and the standard deviation was 5.16. All of them have hold class C1 driving licenses for over 3 years and driven 5,000+ kilometers

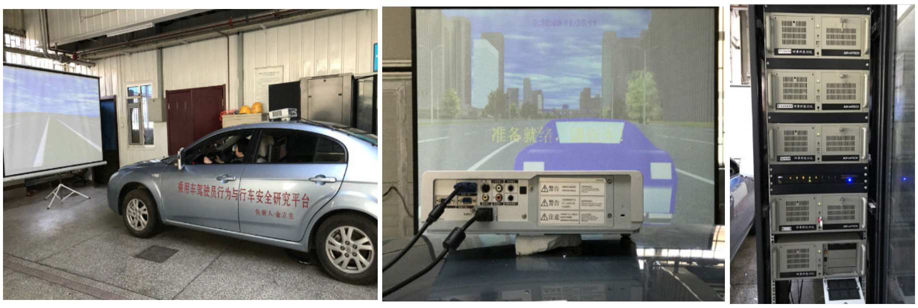

**FIGURE 2.** The driving simulation platform.

**Algorithm 1** Drowsy Driving Detection Model Algorithm

**Input:** Drowsy driving dataset: $G_{n\times p}$; number of information granules: $q$; given parameter of information granulation: $\alpha \in [0, 1]$.

**Output:** Labels of drowsy driving.

1: **for** $i = 1, 2, \ldots, p$ **do**

2: Standardize the dataset $G_{n\times p}$;

3: **end for**

4: **obtain** the standardized dataset $G'_{n\times p}$;

5: Execute factor analysis of the standardized dataset $G'_{n\times p}$ from Eq. (1);

6: **obtain** the common factors: $F_{n\times m}(m \leq p)$;

7: Calculate the composite factor scores of $F_{n\times m}$;

8: **obtain** the series of composite factor scores: $D_{n\times 1}$;

9: Divide the series $D_{n\times 1}$ into $q$ subseries $D_1, D_2, \cdots, D_q$ to create $q$ information granules $\Omega_1, \Omega_2, \cdots, \Omega_q$;

10: **for** $i = 1, 2, \ldots, q$ **do**

11: Perform information granulation to calculate the index $Vol(\Omega_i)$ of the information granule $\Omega_i$ in time window $T_i$ by Eq. (5)-(7);

12: Perform PSO to optimize the time window size from (8);

13: **end for**

14: **obtain** the optimal dynamic time window $T_1^*, T_2^*, \cdots, T_q^*$;

15: Calculate the mean and standard deviation of the composite factors scores in the optimal dynamic time window $T_1^*, T_2^*, \cdots, T_q^*$ and create a training set and a test set;

16: Perform LIBSVM and PSO on the training set by Eqs. (11) -(12), (17);

17: **obtain** the optimal combination $(C, \gamma)$ of penalty coefficient and kernel function parameter;

18: Perform LIBSVM with the optimal combination $(C, \gamma)$ to the test set;

19: **result**: Labels of drowsy driving.

per year. Besides, all the subjects are healthy mentally and physically, without taking any drug that may affect driving in the month preceding our experiments.

### 2) SIMULATION PLATFORM

Our experiments were conducted on the research platform for driver behavior and traffic safety of passenger vehicles, which was developed by Jilin University. As shown in Figure 1, the simulation platform consists of a simulation cab, a projection apparatus and a control cabinet. The platform includes driver touch simulation system, driving environment simulation system, image display system, audio simulation system, vehicle dynamics dynamic simulation system and data acquisition system.

**FIGURE 3.** The photo of the road section.

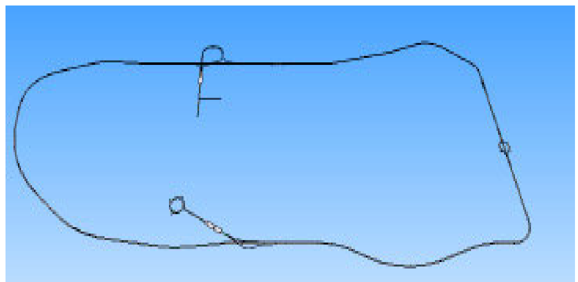

**FIGURE 4.** The alignment of the road section.

For convenience, the Changchun-Siping section of Beijing-Harbin Expressway (Fig. 3 and 4) was selected as the driving scene in our experiments. The 96km-long rectangular road section is essentially a two-way four-lane road, in which each lane is 3.75m-wide and the isolation strip is 3m-wide. The speed limit on vehicles driving on the road section is 80∼120km/h. In addition to the test vehicle, several other vehicles were added, which ran in the same or opposite direction with the test vehicle, aiming to simulate the realistic driving environment.

### 3) EXPERIMENTAL PROCESS

Every driver was required to drive the test vehicle twice, once in normal state and the other in drowsy state. The data of normal driving and drowsy driving were collected into the control group.

**TABLE 1. The characteristic parameters.**

| Type of parameters | Symbols | Name of parameters | Unit |
|---|---|---|---|
| | *Time* | Time | *s* |
| Driver's operating behavior | *SWA* | Steering wheel angle | *rad* |
| | *SWR* | Steering wheel angular speed | *rad/s* |
| | *XA* | Throttle pedal opening | % |
| | *XB* | Brake pedal opening | % |
| | *VSx* | Lateral speed | *m/s* |
| | *VSy* | Longitudinal speed | *m/s* |
| | *ASx* | Lateral acceleration | $m/s^2$ |
| Vehicle's running state | *ASy* | Longitudinal acceleration | $m/s^2$ |
| | *Ya* | Yaw angle | *rad* |
| | *YaR* | Yaw angular speed | *rad/s* |
| | *YaD* | Yaw angular acceleration | $rad/s^2$ |

Based on our circadian clock, drowsy driving tends to occur in three periods: 2:00∼6:00, 11:00∼13:00 and 15:00∼16:00, especially in the first and last periods. The risk of drowsy driving in 15:00∼16:00 is twice that at around 10:00 [54]. Hence, the normal driving experiment was arranged in 9:00∼11:00, while the drowsy driving experiment was scheduled in 14:00∼16:00.

Before the normal driving experiment, the drivers were required to sleep more than 8h a day in the weak leading to the experiment, and intake no alcohol. On the experimental day, the drivers arrived at the test site at 8:30 and had their basic information recorded. Then, the experimenter introduced the simulation platform to the drivers, and gave them 10∼15min to familiarize themselves with the platform. No data was recorded in this period. After that, the normal driving experiment officially started and lasted for 2h.

Before the drowsy driving experiment, the drivers were asked to sleep less, such that they could enter the drowsy state quickly. During the 24h before the experiment, the drivers were required to sleep for less than 5h between 1:00 and 6:00, and take no drug, caffeine or tea. On the experiment day, the drivers first had their basic information recorded. Then, the experimenter introduced the simulation platform to the drivers, and gave them 10∼15min to familiarize themselves with the platform. No data was recorded in this period. After that, the drowsy driving experiment officially started and lasted for 2h.

### B. EXPERIMENTAL RESULTS

As shown in Table 1, the characteristic parameters of driver's operating behavior (e.g. steering wheel angle, steering wheel angular speed, throttle pedal opening, and brake pedal opening) and those of the vehicle's running state (e.g. lateral speed, longitudinal speed, lateral acceleration, longitudinal acceleration, yaw angle, yaw angular speed, and yaw angular acceleration) were collected during the two experiments. Meanwhile, the videos on the foreground and driver's face were taken at the frequency of 10Hz by an external surveillance camera and internal camera.

### C. DROWSY STATE EVALUATION

During the experiments, it is impossible for a driver to evaluate his/her drowsiness in real time. Studies have shown that, if the driver's face video and the road foreground video are evaluated against the Karolinska sleepiness scale (KSS), the results will be almost the same whether the evaluation is done by others or by the driver him/herself [55]. Therefore, the author decided to invite three experts, rather than the driver, to evaluate the driver's drowsiness independently based on the videos against the KSS [56]. The mean KSS score was adopted as the objective score of the drivers' drowsiness. If their scores differed greatly, the three experts were required to reconfirm the original scores. In addition, each driver was asked to evaluate his/her drowsiness before and after each experiment against the KSS.

Drawing on relevant studies [55], a driver was considered as normal if his/her KSS score is smaller than or equal to 3, and drowsy if the latter is greater than or equal to 7. Then, the data labeled of each driver in normal state and those labeled in drowsy state were sorted out for further analysis.

**TABLE 2. Results of the KMO test and bartlett's test.**

| | | |
|---|---|---|
| KMO test | Sampling adequacy | 0.694 |
| Bartlett's test | Approx. Chi-square | 13,311.610 |
| | df | 55 |
| | Sig. | 0.000 |

### D. DATA PREPROCESSING

#### 1) DIMENSIONALITY REDUCTION

For simplicity, the high-dimensional characteristic parameters in Table 1 were processed through factor analysis, using SPSS Statistics 19.0. The eleven parameters, namely, steering wheel angle, steering wheel angular speed, throttle pedal opening, brake pedal opening, lateral speed, longitudinal speed, lateral acceleration, longitudinal acceleration, yaw angle, yaw angular speed and yaw angular acceleration, were denoted as $x_1, x_2, \ldots, x_{11}$, in turn.

Taking the data of driver 1 for example, the data were subjected to Kaiser-Meyer-Olkin (KMO) test and Bartlett's test. The test results (Table 2) show that the data are suitable for factor analysis.

As shown in Fig. 5 and Table 3, there were three common factors with greater-than-one characteristic root, and the first four common factors made a 91.884% cumulative contribution to variance, i.e. the four factors carry about 92% of the original information.

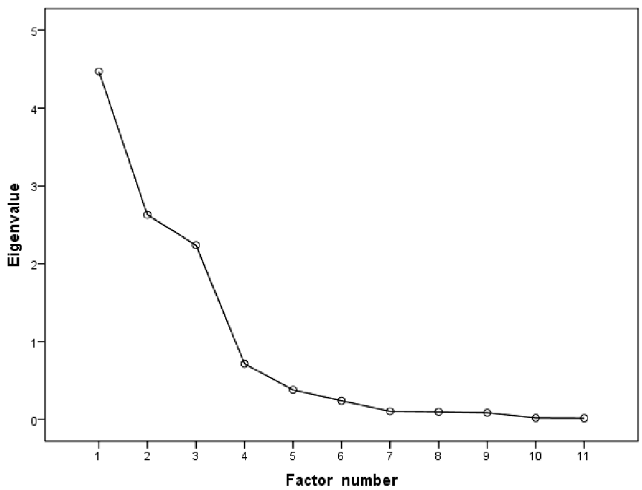


**FIGURE 5.** Eigenvalue-factor number curve.

**TABLE 3.** Eigenvalues, contributions to variance and cumulative contributions to variance of common factors.

| Factors | Eigenvalue | Contribution to variance (%) | Cumulative Contribution to variance (%) |
|---|---|---|---|
| $F_1$ | 4.308 | 39.168 | 39.168 |
| $F_2$ | 2.739 | 24.901 | 64.069 |
| $F_3$ | 2.166 | 19.695 | 83.764 |
| $F_4$ | 0.893 | 8.120 | 91.884 |

**TABLE 4.** Rotated factor matrix.

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $x_1$ | 0.227 | 0.857 | 0.351 | 0.180 |
| $x_2$ | | | 0.966 | |
| $x_3$ | 0.698 | 0.158 | | 0.605 |
| $x_4$ | 0.973 | | | |
| $x_5$ | 0.955 | 0.102 | | 0.178 |
| $x_6$ | | -0.934 | 0.234 | |
| $x_7$ | 0.205 | 0.181 | | 0.938 |
| $x_8$ | | -0.217 | -0.865 | |
| $x_9$ | 0.982 | | | |
| $x_{10}$ | | 0.968 | 0.100 | 0.156 |
| $x_{11}$ | | -0.104 | 0.893 | |

As shown in Table 1, $F_1$ had significant positive correlations with $x_3, x_4, x_5$ and $x_9$; $F_2$ had significant positive correlations with $x_1$ and $x_{10}$, and a significant negative correlation with $x_6$; $F_3$ had significant positive correlations with $x_2$ and $x_{11}$, and a significant negative correlation with $x_8$; $F_4$ had significant positive correlations with $x_3$ and $x_7$.

According to Table 5, the four common factors can be expressed as:

$$\begin{aligned} F_1 = &\, 0.007Zx_1 + 0.002Zx_2 + 0.100Zx_3 + 0.343Zx_4 \\ &+ 0.309Zx_5 + 0.021Zx_6 - 0.172Zx_7 + 0.346Zx_9 \\ &- 0.056Zx_{10} + 0.001Zx_{11}, \\ F_2 = &\, 0.311Zx_1 - 0.035Zx_2 - 0.056Zx_3 - 0.018Zx_4 \\ &- 0.022Zx_5 - 0.399Zx_6 + 0.080Zx_7 - 0.064Zx_8 \\ &- 0.021Zx_9 + 0.378Zx_{10} - 0.085Zx_{11}, \end{aligned}$$

**TABLE 5.** Factor score coefficient matrix.

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $x_1$ | 0.007 | 0.311 | 0.099 | -0.017 |
| $x_2$ | 0.002 | -0.035 | 0.366 | -0.003 |
| $x_3$ | 0.100 | -0.056 | -0.010 | 0.400 |
| $x_4$ | 0.343 | -0.018 | 0.003 | -0.185 |
| $x_5$ | 0.309 | -0.022 | -0.003 | -0.084 |
| $x_6$ | 0.021 | -0.399 | 0.125 | 0.145 |
| $x_7$ | -0.172 | -0.080 | -0.013 | 0.857 |
| $x_8$ | 0.000 | -0.064 | -0.320 | 0.074 |
| $x_9$ | 0.346 | -0.021 | -0.000 | -0.186 |
| $x_{10}$ | -0.056 | 0.378 | 0.001 | 0.013 |
| $x_{11}$ | 0.001 | -0.085 | 0.342 | 0.044 |

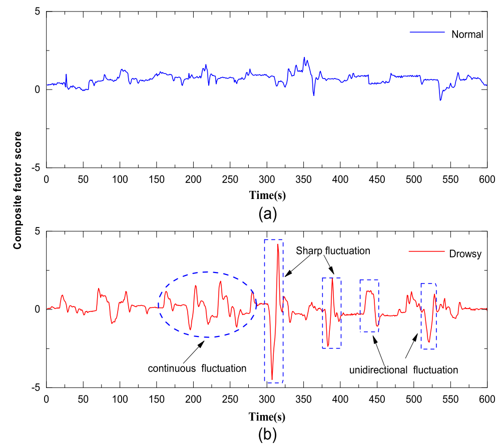


**FIGURE 6.** Comparison of composite factor scores in (a) normal state and (b) drowsy state.

$$\begin{aligned} F_3 = &\, 0.099Zx_1 + 0.369Zx_2 - 0.010Zx_3 + 0.003Zx_4 \\ &- 0.003Zx_5 + 0.125Zx_6 - 0.013Zx_7 - 0.320Zx_8 \\ &- 0.001Zx_{10} + 0.342Zx_{11}, \\ F_4 = &\, -0.017Zx_1 - 0.003Zx_2 + 0.400Zx_3 - 0.185Zx_4 \\ &- 0.084Zx_5 + 0.145Zx_6 + 0.857Zx_7 + 0.074Zx_8 \\ &- 0.186Zx_9 - 0.013Zx_{10} + 0.044Zx_{11}. \end{aligned}$$

where, $Zx_1, Zx_2, \cdots, Zx_{11}$ are the normalized variables of $x_1, x_2, \cdots, x_{11}$. Then, the composite factor score of driver 1 in normal or drowsy state can be described as:

$$F = 0.39F_1 + 0.25F_2 + 0.20F_3 + 0.08F_4.$$

The data of the other drivers were processed in the same manner, resulting in the composite factor scores of these drivers in the normal and drowsy states. Then, the composite factor scores of the two states were compared (Fig. 6).

From Fig. 6, it can be seen that, under normal driving, the composite factor score was basically stable, with a slight fluctuation. Under drowsy state, the composite factor score was very unstable, featuring violent fluctuations.

Furthermore, the composite factor scores in the normal and drowsy states were subjected to a one-way analysis of variance (ANOVA) on SPSS Statistics 19.0, with the significance level of 0.05. The analysis results (Table 6), $F = 19.722$ and

**TABLE 6.** ANOVA results.

| | Sum of squares | df | Mean square | F | Sig. |
|---|---|---|---|---|---|
| Between group variation | 14.780 | 1 | 14.780 | 19.722 | 0.000 |
| Intra-group variance | 5218.820 | 6964 | 0.749 | | |
| Total | 5233.600 | 6965 | | | |

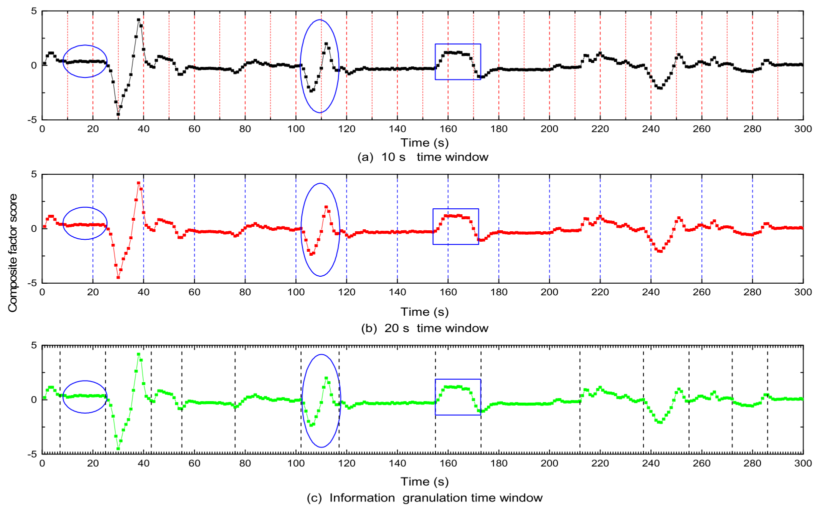


**FIGURE 7.** The variations of the composite factor scores of a driver in different time windows.

$p = 0.000 < 0.05$, show that the composite factor scores can characterize the features of drowsy driving.

### 2) OPTIMIZATION OF DYNAMIC TIME WINDOW

The optimal time window of characteristic parameters was searched for through information granulation of time series. Specifically, the time series of the composite factor scores of each driver was divided into 5min-long segments. The data shorter than 5min were discarded. In this way, the author obtained 162 drowsing state segments and 214 normal state segments. Then, each segment was processed into 15 information granules. The particle swarm optimization (PSO) was introduced to optimize the dynamic time window of common factors in each segment. Fig. 7 presents the variations of the composite factor scores of a driver in different time windows in 300s. Obviously, the dynamic time window is suitable for partitioning of data states.

## *E. MODEL CONSTRUCTION AND RESULTS ANALYSIS*

The mean and standard deviation of the composite factors scores in the optimal dynamic window were fused. Then, the characteristic parameters of normal and drowsy states were extracted from the fused data, producing 3,210 normal samples and 2,430 drowsy samples. Then, the LIBSVM toolbox was employed to set up a drowsy driving detection model.

### 1) MODEL TRAINING

The model training, essentially the selection of kernel function, aims to optimize the internal parameters using the training samples. In this paper, the radial basis function (RBF) is selected as the kernel function:

$$k(x_i, y_j) = e^{-\gamma|x_i,y_j|^2} \tag{19}$$

Then, the model was trained to optimize the penalty coefficient $C$ and kernel function parameter $\gamma$. Firstly, 2,000 normal samples and 1,600 drowsy samples were selected randomly and grouped into the training set. Taking accuracy as the training objective, the PSO was conducted to find the optimal combination $(C, \gamma)$ of penalty coefficient and kernel function parameter. The results show that the optimal value of $C$ and $\gamma$ were respectively 0.5 and 0.3.

### 2) MODEL TESTING AND RESULTS

The established model was tested with the test set, which contains the remaining 1,210 normal samples and 830 drowsy samples. To verify its effectiveness, our model was compared with several other models, which are also based on the data fused between the mean and standard deviation of composite factor scores. The main difference between our model and the contrastive models lies in that: our model adopts the optimal dynamic time window with LIBSVM, while those models adopt a fixed time window (20s) generated by LIBSVM, decision tree [57], backpropagation [58] and fuzzy c-means (FCM) clustering [59], respectively. Table 7 compares the detection accuracies of the models, including the true positive rate (TPR), true negative rate (TNR) and the false positive rate (FPR). The TRP refers to the detection accuracy of drowsy state, the TNR refers to the detection accuracy of normal state, and the FRP refers to the proportion of normal states that are wrongly determined as drowsy states. The receiver

**TABLE 7. The detection accuracies of the two models.**

| Method | Actual state of the driver | Detection result | | Detection accuracy | | | |
|---|---|---|---|---|---|---|---|
| | | Drowsy | Normal | TPR | TNR | FPR | Overall accuracy |
| Our model | Drowsy | 788 | 42 | 94.9% | 80.7% | 19.3% | 86.47% |
| | Normal | 234 | 976 | | | | |
| Contrastive model 1 | Drowsy | 775 | 55 | 93.3% | 73.9% | 26.1% | 81.81% |
| | Normal | 316 | 894 | | | | |
| Contrastive model 2 | Drowsy | 768 | 62 | 92.5% | 72.4% | 27.6% | 80.58% |
| | Normal | 334 | 876 | | | | |
| Contrastive model 3 | Drowsy | 762 | 68 | 91.8% | 67.8% | 32.2% | 77.55% |
| | Normal | 390 | 820 | | | | |
| Contrastive model 4 | Drowsy | 754 | 76 | 90.8% | 66.4% | 33.6% | 76.32% |
| | Normal | 407 | 803 | | | | |

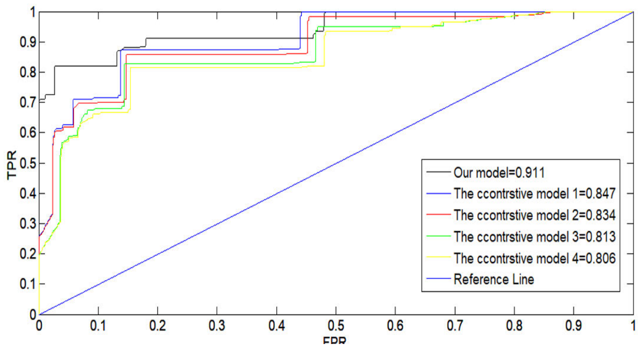


**FIGURE 8. The ROC curves of the models.**

operating characteristic (ROC) curves and the areas under the curve (AUC) of the models are given in Fig. 8.

## IV. DISCUSSION

This paper generally detects drowsy driving based on the fused data on the characteristic parameters of driver's operating behavior and the vehicle's running state. The four basic steps of our research are discussed below.

### A. DIMENSIONALITY REDUCTION OF CHARACTERISTIC PARAMETERS

The factor analysis was conducted to reduce the dimensions of the characteristic parameters of driver's operating behavior and the vehicle's running state. Through the analysis, the author established the time series of the composite factor scores. The advantage of this strategy lies in the identification of main influencing factors and their impacts through the analysis, and the ability to determine the composite factor scores. As shown in Table 3, the 11 characteristic parameters were reduced to 4 common factors, while preserving 92% of the original information. The results in Fig. 6 and Table 6 show that the composite factor scores are significant indices of drowsy driving detection.

### B. OPTIMIZATION OF DYNAMIC TIME WINDOW

In previous studies, the characteristic parameters are usually fixed in the same time window, without considering the dynamic features of the time window. As shown in Fig. 7(a) and (b), when the time window was fixed at 10 or 20s, the changes in parameter data could not be tracked timely, making it impossible to interpret the data in a rational manner. Neither was the results satisfactory when the time window was adjusted to 15 or 30s. To solve the problem, this paper adopts the information granulation of time series, splitting the time series of common factors into information granules. Next, the PSO was performed to optimize the dynamic time window. Fig. 7(c) shows that the time window size obtained by granulation was not fixed but varied with the data trend. Compared with fixed time windows, the dynamic time window leads to high data fitness and significant indices in each sub-time window.

### C. INFORMATION FUSION

In each sub-time window, the mean and standard deviation of composite factor scores were computed and fused into one data. The characteristic parameters of the driver's operating behavior were combined with those of the vehicle's running state to enhance the accuracy of drowsy driving detection.

### D. CONSTRUCTION OF DETECTION MODEL

Based on the fused information, the normal samples and drowsy samples were determined, respectively. Then, the samples were divided into a training set and a test set. Our drowsy driving detection model was trained and tested by LIBSVM. After that, our model, which has a dynamic time window, was compared with models with fixed time window. Table 7 and Fig. 8 demonstrate that our model outperformed the contrastive models with an accuracy of 86.47% and an AUC of 0.911.

## V. CONCLUSION

For accurate detection of drowsy driving, this paper put forward a novel model based on data fusion and information granulation. The drowsing driving mode was identified on a simulation platform for driver's operating behavior and the vehicle's running state. In our drowsy driving experiments, 16 drivers were recruited to drive the test vehicle twice, once in normal state and the other in drowsy state, and 11 indicative parameters of both normal and drowsy states were collected as evaluation indices. The factor analysis was applied to reduce the dimensionality and determine the composite factor scores of the indicative parameters. Through the analysis, the indicative parameters were reduced to 4 common factors, while preserving 92% of the original information. In addition, the variations of the composite factor scores were compared between the fixed time window and the dynamic time window, revealing that the latter can track and interpret the data variations timelier and more rationally. Therefore, the time series of common factors was split into information granules with the strategy called the information granulation of time series, and the optimal dynamic time window was obtained by the PSO. Finally, the drowsy driving detection model was trained and tested by LIBSVM based on the fused information of the mean and standard deviation of composite factor scores in each sub-time window. Experimental results showed that our model outperformed the contrastive models with an accuracy of 86.47% and an AUC of 0.911.

In this paper, drowsy driving detection is investigated under the condition that the driver is not distracted by other factors. In the future, more attention will be paid to the drowsy driving detection when the driver is in distracted states. After all, in real-world scenarios, the driver is often distracted by factors like chatting, listening to the music and making a phone call.

## REFERENCES


[1] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, and D. J. Ramsey, ''The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data,'' Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 810 594, Jan. 2006.

[2] P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, and H. P. A. Van Dongen, ''Efficient driver drowsiness detection at moderate levels of drowsiness,'' *Accident Anal. Prevention*, vol. 50, pp. 341–350, Jan. 2013.

[3] X. Zhao, S. Xu, J. Rong, and X. Zhang, ''Discriminating threshold of driving fatigue based on the electroencephalography sample entropy by receiver operating characteristic curve analysis,'' *J. Southwest Jiaotong Univ.*, vol. 48, no. 1, pp. 178–183, 2013.

[4] J. S. Higgins, J. Michael, R. Austin, T. Åkerstedt, and P. Hans, ''Asleep at the wheel—The road to addressing drowsy driving,'' *Sleep*, vol. 40, no. 2, pp. 1–9, 2017.

[5] P. Zheng, ''Study situation and developing trend on detecting and evaluating technique of motor driver fatigue,'' *J. China Agric. Univ.*, vol. 6, no. 6, pp. 101–105, 2001.

[6] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, ''Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm,'' *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, Sep. 2011.

[7] V. M. Y. Mervyn, X. Li, K. Shen, and E. P. V. Wilder-Smith, ''Can SVM be used for automatic EEG detection of drowsiness during car driving?'' *Saf. Sci.*, vol. 47, no. 1, pp. 115–124, 2009.

[8] C.-T. Lin, C.-J. Cheng, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang, ''A real-time wireless brain-computer interface system for drowsiness detection,'' *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 4, pp. 214–222, Apr. 2015.

[9] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, ''Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system,'' *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 9, pp. 2044–2055, Sep. 2012.

[10] K. Fujiwara, E. Abe, K. Kamata, C. Nakayama, Y. Suzuki, T. Yamakawa, T. Hiraoka, M. Kano, Y. Sumi, F. Masuda, M. Matsuo, and H. Kadotani, ''Heart rate variability-based driver drowsiness detection and its validation with EEG,'' *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1769–1778, Nov. 2018.

[11] H. Wang, C. Wu, T. Li, Y. He, P. Chen, and A. Bezerianos, ''Driving fatigue classification based on fusion entropy analysis combining EOG and EEG,'' *IEEE Access*, vol. 7, pp. 61975–61986, 2019.

[12] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, and S. C. Jun, ''Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data,'' *Front. Hum. Neurosci.*, vol. 10, pp. 1–14, May 2016.

[13] F. Guede-Fernández, M. Fernández-Chimeno, J. Ramos-Castro, and M. A. García-González, ''Driver drowsiness detection based on respiratory signal analysis,'' *IEEE Access*, vol. 7, pp. 81826–81838, 2019.

[14] N. T. B. Pasaribu, A. Prijono, R. Ratnadewi, R. P. Adhie, and J. Felix, ''Drowsiness detection according to the number of blinking eyes specified from eye aspect ratio value modification,'' in *Proc. 1st Int. Conf. Life, Innov., Change Knowl.*, vol. 203, 2019, pp. 171–174.

[15] A. M. Rumagit, I. A. Akbar, M. Utsunomiya, T. Morie, and T. Igasaki, ''Gazing as actual parameter for drowsiness assessment in driving simulators,'' *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 13, no. 1, pp. 170–178, 2019.

[16] W. W. Wierwille and L. A. Ellsworth, ''Evaluation of driver drowsiness by trained raters,'' *Accident Anal. Prevention*, vol. 26, pp. 571–581, Oct. 1994.

[17] Q. Wu, B. X. Sun, B. Xie, and J. Zhao, ''A PERCLOS-based driver fatigue recognition application for smart vehicle space,'' in *Proc. 3rd Int. Symp. Inf. Process. (ISIP)*, Oct. 2010, pp. 437–441.

[18] J. J. Yan, H.-H. Kuo, Y.-F. Lin, and T.-L. Liao, ''Real-time driver drowsiness detection system based on PERCLOS and grayscale image processing,'' in *Proc. IEEE Int. Symp. Comput. Consum. Control. (IS3C)*, Jul. 2016, pp. 243–246.

[19] G. Soares, D. De Lima, and A. M. Neto, ''A mobile application for driver's drowsiness monitoring based on PERCLOS estimation,'' *IEEE Latin Amer. Trans.*, vol. 17, no. 2, pp. 193–202, Oct. 2019.

[20] M. Saradadevi and P. Bajaj, ''Driver fatigue detection using mouth and yawning analysis,'' *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 6, pp. 183–188, 2008.

[21] S. Abtahi, B. Hariri, and S. Shirmohammadi, ''Driver drowsiness monitoring based on yawning detection,'' in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2011, pp. 1–4.

[22] Z. Jie, M. Mahmoud, Q. Stafford-Fraser, P. Robinson, E. Dias, and L. Skrypchuk, ''Analysis of yawning behaviour in spontaneous expressions of drowsy drivers,'' in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 571–576.

[23] A. D. McDonald, C. Schwarz, J. D. Lee, and T. L. Brown, ''Real-time detection of drowsiness related lane departures using steering wheel angle,'' in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 56, no. 1. Los Angeles, CA, USA: Sage, 2012, pp. 2201–2205.

[24] Z. Li, S. E. Li, R. Li, B. Cheng, and J. Shi, ''Online detection of driver fatigue using steering wheel angles for real driving conditions,'' *Sensors*, vol. 17, no. 3, pp. 1–12, 2017.

[25] S. Arefnezhad, S. Samiee, A. Eichberger, and A. Nahvi, ''Driver drowsiness detection based on steering wheel data applying adaptive neuro-fuzzy feature selection,'' *Sensors*, vol. 19, no. 4, p. 943, 2019.

[26] M. Chai, S.-W. Li, W.-C. Sun, M.-Z. Guo, and M.-Y. Huang, ''Drowsiness monitoring based on steering wheel status,'' *Transp. Res. D, Transp. Environ.*, vol. 66, pp. 95–103, Jan. 2019.

[27] C. Schwarz, J. Gaspar, T. Miller, and R. Yousefian, ''The detection of drowsiness using a driver monitoring system,'' *Traffic Injury Prevention*, vol. 20, no. 1, pp. S157–S161, 2019.

[28] Y. Xiao, C. Song, and Y. Liu, ''Social hotspot propagation dynamics model based on multidimensional attributes and evolutionary games,'' *Commun. Nonlinear Sci. Numer. Simul.*, vol. 67, pp. 13–25, Feb. 2018.

[29] Y. Xiao, D. Chen, S. Wei, Q. Li, H. Wang, and M. Xu, ''Rumor propagation dynamic model based on evolutionary game and anti-rumor,'' *Nonlinear Dyn.*, vol. 95, no. 1, pp. 523–539, 2019.

[30] L. Jin, B. Guo, 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, pp. 65772–65782, 2018.

[31] B. Guo, L. Jin, D. Sun, J. Shi, and F. Wang, ''Establishment of the characteristic evaluation index system of secondary task driving and analyzing its importance,'' *Transp. Res. F, Traffic Psychol. Behav.*, vol. 64, pp. 308–317, Jul. 2019.

[32] K. J. Raman, A. Azman, V. Arumugam, S. Z. Ibrahim, S. Yogarayan, M. F. A. Abdullah, S. F. A. Razak, A. H. M. Amin, and K. Sonaimuthu, ''Fatigue monitoring based on Yawning and head movement,'' in *Proc. 6th Int. Conf. Inf. Commun. Technol. (ICoICT)*, May 2018, pp. 343–347.

[33] K. Y. Li, *Research on Drowsiness Driving Detection Considering Driving Style*. Jilin, China: Jilin Univ., 2017.

[34] C. Wang, Y. Tian, and H. Jia, "Driving fatigue detection based on feature fusion of information entropy," *J. Comput. Methods Sci. Eng.*, vol. 18, no. 4, pp. 977–988, 2018.

[35] X. Li, L. Hong, J. Wang, and X. Liu, "Fatigue driving detection model based on multi-feature fusion and semi-supervised active learning," *IET Intell. Transp. Syst.*, vol. 13, no. 9, pp. 1401–1409, Sep. 2019.

[36] W. Liu, J. Qian, Z. Yao, X. Jiao, and J. Pan, "Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection," *Future Internet*, vol. 11, no. 5, pp. 1–13, 2019.

[37] A. Eskandarian, R. Sayed, P. Delaigue, A. Mortazavi, and J. Blum, "Advanced driver fatigue research," U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. FMCSA-RRR-07-001, 2007.

[38] J. Zhang, Z. X. Zhu, Z. H. Song, and E. R. Mao, "Research on information fusion on evaluation of driver fatigue," in *Proc. Int. Symp. Comput. Sci. Comput. Technol. (ISCSCT)*, vol. 2, 2008, pp. 151–155.

[39] Q. N. Niu, *Research on Driver Fatigue Detection Based on Hybrid Measures*. Jilin, China: Jilin Univ., 2014.

[40] M. S. Wang, N. T. Jeong, K. S. Kim, S. B. Choi, S. M. Yang, S. H. You, J. H. Lee, and M. W. Suh, "Drowsy behavior detection based on driving information," *Int. J. Automot. Technol.*, vol. 17, no. 1, pp. 165–173, Feb. 2016.

[41] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets Syst.*, vol. 90, no. 2, pp. 111–127, 1997.

[42] W. Pedrycz and F. Gomide, "Fuzzy systems engineering: Toward human-centric computing," in *Fuzzy Systems Engineering: Toward Human-Centric Computing*. Hoboken, NJ, USA: Wiley, 2007, pp. 1–526.

[43] W. Wang, W. Pedrycz, and X. Liu, "Time series long-term forecasting model based on information granules and fuzzy clustering," *Eng. Appl. Artif. Intell.*, vol. 41, pp. 17–24, May 2015.

[44] F. J. Cabrerizo, W. Pedrycz, I. J. Pérez, S. Alonso, and E. Herrera-Viedma, "Group decision making in linguistic contexts: An information granulation approach," *Procedia Comput. Sci.*, vol. 91, pp. 715–724, Jul. 2016.

[45] W. Lu, W. Pedrycz, J. Yang, and X. Liu, "Granular fuzzy modeling guided through the synergy of granulating output space and clustering input subspaces," *IEEE Trans. Cybern.*, to be published.

[46] M. Xu, J. Wu, M. Liu, Y. Xiao, H. Wang, and D. Hu, "Discovery of critical nodes in road networks through mining from vehicle trajectories," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 583–593, Feb. 2019.

[47] C. Spearman, "'General Intelligence,' objectively determined and measured," *Amer. J. Psychol.*, vol. 15, no. 2, pp. 201–292, 1904.

[48] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[49] R. Eberhart and J. Kennedy, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.

[50] Z. Pawlak, "Granularity of knowledge, indiscernibility and rough sets," in *Proc. IEEE Int. Conf. Fuzzy Syst., IEEE World Congr. Comput. Intell.*, vol. 1, May 1998, pp. 106–110.

[51] Z. Ling and Z. Bo, "Theory of fuzzy quotient space (methods of fuzzy granular computing)," *J. Softw.*, vol. 14, no. 4, pp. 770–776, 2003.

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

[53] T. Özkan and T. Lajunen, "What causes the differences in driving between young men and women? The effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 9, no. 4, pp. 269–277, 2006.

[54] J. A. Horne and L. A. Reyner, "Driver sleepiness," *J. Sleep Res.*, vol. 4., pp. 23–29, Dec. 1995.

[55] M. Ingre, T. Åkerstedt, B. Peters, A. Anund, and G. Kecklund, "Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences," *J. Sleep Res.*, vol. 15, no. 1, pp. 47–53, 2006.

[56] T. Akerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *Int. J. Neurosci.*, vol. 52, nos. 1–2, pp. 29–37, Jul. 1990.

[57] E. B. Hunt, J. Marin, and P. J. Stone, *Experiments in Induction*. New York, NY, USA: Academic, 1966.

[58] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[59] J. C. Bezdek, W. Full, and R. Ehrlich, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.

**YAN WANG** received the B.S. degree in mathematics education from Jilin Normal University, Siping, China, in 2002, and the M.S. degree from Northeast University, Shenyang, China, in 2008. He is currently pursuing the Ph.D. degree with the Department of Vehicle Operation Engineering, Transportation College, Jilin University. His research interests include vehicle safety and intelligent vehicle navigation technology, and driver behavior analysis.

**LISHENG JIN** received the B.S. degree in construction machinery, the M.S. degree in mechanical design and theory, and the Ph.D. degree in mechatronic engineering from Jilin University, Changchun, China, in 1997, 2000, and 2003, respectively. He is currently a Professor with Jilin University. He has achieved valuable research results and published nearly 100 academic articles. His research interests include vehicle safety and intelligent vehicle navigation technology, vehicle ergonomics, and driver behavior analysis.

**KEYONG LI** received the B.S. degree in vehicle engineering from Northeast Forestry University, Harbin, China, in 2012, and the Ph.D. degree in vehicle operation engineering from Jilin University, Changchun, China, in 2017. He is currently a Research Engineer with Shanghai Huawei Technologies Company, Ltd. His research focuses on HMI, including vehicle ergonomics, driver behavior analysis, and vehicle safety.

**BAICANG GUO** received the B.S. degree in electrical engineering and its automation from Northeast Forestry University, Harbin, China, in 2014, and the M.S. degree from Jilin University, Changchun, China, in 2018. He is currently pursuing the Ph.D. degree with the Department of Vehicle Operation Engineering, Transportation College, Jilin University. His research interests include vehicle safety and intelligent vehicle navigation technology, and driver behavior analysis.

**YI ZHENG** received the B.S. degree in automation from Jilin University, Changchun, China, in 2008. He is currently pursuing the Ph.D. degree with the Department of Vehicle Operation Engineering, Transportation College, Jilin University. His research interests include ADAS, intelligent connected vehicles, and vehicle safety.

**JIAN SHI** received the B.S. degree in electrical engineering and its automation from the Changchun University of Technology, Jilin, China, in 2017. He is currently pursuing the M.S. degree in control science and engineering with the College of Communication, Jilin University. His research interests include machine learning, deep learning, computer vision, and intelligent vehicle technology.