

## Research Article

# Extraction of Optimal Measurements for Drowsy Driving Detection considering Driver Fingerprinting Differences

Yifan Sun , Chaozhong Wu , Hui Zhang , Yijun Zhang , Shaopeng Li ,  
and Hongxia Feng 

*Intelligent Transportation Systems Research Center, Wuhan University of Technology, 1178 Hepingdadao Street, Wuchang District, Wuhan 430063, Hubei, China*

Correspondence should be addressed to Hui Zhang; zhanghuiits@whut.edu.cn

Received 21 January 2021; Revised 14 June 2021; Accepted 11 August 2021; Published 31 August 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yifan Sun 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.

Contributions of measurements for detecting drowsy driving are determined by calculation parameters, which are directly related to the accuracy of drowsiness detection. The previous studies utilized the same Unified Calculation Parameters (UCPs) to compute each driver's measurements. However, since each driver has unique driving behavior characteristics, namely, driver fingerprinting, Individual Drivers' Best Calculation Parameters (IDBCPs) making measurements more discriminative for drowsiness are various. Regardless of the difference in driver fingerprinting among the drivers being tested, using UCPs instead of IDBCPs to compute measurements will limit the drowsiness-detection performance of the measurements and reduce drowsiness-detection accuracies at the individual driver level. Thus, this paper proposed a model to optimize calculation parameters of individual driver's measurements and to extract individual driver's measurements that effectively distinguish drowsy driving. Through real vehicle experiments, we collected naturalistic driving data and subjective drowsy levels evaluated by the Karolinska Sleepiness Scale. Eight nonintrusive drowsiness-related measurements were calculated by double-layer sliding time windows. In the proposed model, we firstly applied the Wilcoxon test to analyze differences between measurements of the awake state and drowsy state, and constructed the fitness function reflecting the relationship between the calculation parameters and measurement's drowsiness-detection performance. Secondly, the genetic algorithms were used to optimize fitness functions to obtain measured IDBCPs. Finally, we selected measurements calculated by IDBCPs that can distinguish drowsy driving to constitute individual drivers' optimal drowsiness-detection measurement set. To verify the advantages of IDBCPs, the measurements calculated by UCPs and IDBCPs were, respectively, used to build driver-specific drowsiness-detection models: DF\_U and DF\_I based on the Fisher discriminant algorithm. The mean drowsiness-detection accuracies of DF\_U and DF\_I were, respectively, 85.25% and 91.06%. It indicated that IDBCPs could enhance measurements' drowsiness-detection performance and improve the drowsiness-detection accuracies. This paper contributed to the establishment of personalized drowsiness-detection models considering driver fingerprinting differences.

## 1. Introduction

Drowsy driving is a typically dangerous driving behavior, which seriously threatens traffic safety and causes substantial financial costs to individuals and society [1–3]. Drowsy driving is a common phenomenon that results from multiple underlying reasons, including prolonged driving hours and sleep deprivation [3, 4]. A review of previous studies indicated that there was an obvious association between drowsy driving and the risk of traffic accidents [4, 5]. The studies pointed out that 7% of all

accidents and 16.5% of all road casualties were related to drowsy driving [6]. And, in the European Union, about 20% of commercial transport crashes were caused by drowsy drivers [7]. It was found that the risk of a near-crash event was significantly increased when drivers were drowsy after a night shift [8]. Besides, it is a problem to judge whether a traffic accident is caused by drowsiness [9]. Unlike drunk driving, many accidents related to drowsiness have not been reported due to the lack of objective criteria for judging drowsiness occurrence, and the actual hazards of drowsiness driving may be more serious [10].

Therefore, it is of great significance to study the accurate real-time anti-drowsiness warning system.

Drowsiness-detection models are at the core of the anti-drowsiness warning systems, which are mainly divided into two categories: intrusive and nonintrusive [11, 12]. The intrusive drowsiness-detection models rely on contact devices to collect physiological data, such as electroencephalograms [13], electromyography [14], etc. For example, researchers [15] obtained the signal of heart rate variability (HRV) from surface electrocardiogram and trained drowsiness detector using 7 features derived from HRV, and the positive predictive value reached 0.96. Although the drowsiness-detection using physiological signal is more accurate, the intrusiveness of collecting data limits the practicability [14]. The nonintrusive drowsiness-detection models use noncontact sensors to collect data nonintrusively to detect drowsy driving, such as steering wheel angle, lane position, eye movement data, etc. [14, 16]. The nonintrusive drowsiness-detection models are more practical because they have little interference with drivers and fewer restrictions [14]. So, nonintrusive drowsiness-detection methods have always been a hotspot in drowsy driving research [12, 17]. As reported in much literature, the researchers established various drowsiness-detection models based on nonintrusive measurements, such as lane position derived measurements [18], steering wheel angle derived measurements [17], and the ocular movement derived measurements [19] and obtained relatively accurate drowsiness-detection results. Thus, this paper concentrated on nonintrusive measurements and models of drowsiness detection.

The sensitivity of nonintrusive measurements to drowsiness depends on the time window setting method and calculation parameters of measurements, which greatly determines the accuracy of drowsiness-detection [18, 19]. Many studies concerning measurements and detection methods of drowsy driving have been conducted, which can be found in some review literature [11, 12, 14, 20–22]. This paper focuses on the optimization of calculation parameters of nonintrusive drowsiness measurements. Therefore, we summarize some literature that typically uses nonintrusive measurements to detect drowsiness and gives specific computational parameters. Table 1 is about the apparatus, time window setting, and parameters for nonintrusive measurements calculation in partial drowsy driving studies. It can be seen that time window setting and calculation parameters of nonintrusive measurements in studies are inconsistent, which indicates that these factors affect research outcomes of drowsy driving. Therefore, the calculation parameters of the measurements need to be optimized to improve the drowsiness-detection performance of nonintrusive measurements.

Most of the drowsy driving studies calculated the measurements of each driver by the same Unified Calculation Parameters (UCPs), without considering individual differences in measurements' calculation parameters, which are results of driver fingerprinting differences among drivers. It has been discovered that each person not only has unique physical characteristics but also behavioral characteristics [31]. For example, every person has his or her behavioral characteristics in shopping, writing, reading, etc. Driver

fingerprinting is each driver's unique driving behavior characteristics during the driving duration [32], which are reflected in the steering wheel angle, speed, eye movement, and other driving behavior data [33, 34]. The driver fingerprinting is stable and unique [33–35]. So, several researchers have tried to describe driver fingerprinting to identify the drivers. Finker et al. [35] collected multiple CAN-bus signals by field experiments, and cepstral feature extraction techniques were used to analyze brake pedal signals and gas pedal signals. They used artificial neural networks to learn driver fingerprinting characteristics and performed driver-identification. The identification accuracy yielded 84.6%. Xun et al. [36] used actual vehicles to collect naturalistic driving behavior data, including speed, steering wheel angle, accelerator pedal signal, etc. They built a model to achieve driver fingerprinting based on a convolutional neural network and support vector domain description and used driver fingerprinting to accurately identify drivers. The above studies utilized the differences in driver fingerprinting among drivers to accurately identify drivers, which proves that drivers have stable driving behavior characteristics that are different from others. And, the differences in driver fingerprinting were also the reason for the individual differences in some aspects of drowsy driving studies, such as the calculation parameters, data distribution, and drowsiness-detection threshold of drowsiness-related measurements [37–39].

Many scholars have noticed individual differences caused by driver fingerprinting differences in drowsy driving detection research [18, 38, 40, 41]. Most drowsiness-detection models were generalized models that were trained by blending measurements of all drivers, and these models detected the drowsiness of each driver utilizing the same drowsiness threshold [14]. The differences in driver fingerprinting were the important reasons for the low reliability of the nonintrusive generalized drowsiness-detection models [39, 42, 43]. It was frequently mentioned that, due to the driver fingerprinting differences, drivers differed considerably in data distribution of measurements such as SDLP [39], the standard deviation of steering wheel movement (SDSWM) [38], PERCLOS [39], eye closure duration [43], etc. For instance, Inger et al. [37] analyzed the correlation between self-reported drowsiness level and blink duration and SDLP. The results indicated that there were obvious differences among drivers in the distribution of blink duration and SDLP under various levels of drowsiness. They also pointed out that the drowsiness-detection at the individual driver level suffered from systematic errors if the drowsy driving was detected by the same drowsiness-detection threshold for all drivers. Moreover, researchers explored reasons for differences in driver fingerprinting that affected the drowsiness-identification. Silveira, et al. adopted features derived from electrocardiographs and eye movement signals to build the classifiers of subject-dependent driver sleepiness. They found that physiological signals presented obvious individual characteristics, while subject-independent drowsiness classifiers that ignored individual differences performed worse [44]. Similarly, Persson, et al. used HRV measurements to build a drowsiness classifier based on machine learning and pointed out that the accuracy

TABLE 1: Summary of information in drowsy driving studies.

References	Apparatus	Time window setting	Nonintrusive measurements	Calculation parameters
Li et al. [23]	RV	SSTW	ApEn	Time window is 60 s
Wakita et al. [24]	DS	SPTW	VCF and DCF	Time window is 0.6 s
Rossi et al. [25]	DS	SPTW	Mean and standard deviation of steering error and SDLP	The time window is 60 s
Henni et al. [26]	RV	SPTW	Eye closure duration, mean of speed, and steering wheel movement	The time window is 30 s
Feng et al. [27]	DS	SPTW	PERCLOS, average eye-opening level, PNS, and SDSWA	The time window is 30 s
McDonald et al. [28]	DS	SPTW	Steering wheel angle	The time window is 60 s
Zandi et al. [19]	DS	SSTW	PERCLOS, saccade rate, and blinking rate	The time window ranges from 30 s to 60 s
Cheng et al. [29]	DS	SPTW	PERCLOS blinking duration, and blinking rate	Time window is 60 s
Zhang et al. [30]	RV	DPTW	SDSWA and PNS	The first-layer time window is 60 s, the second-layer time window is 15 s

RV = real vehicle, DS = driving simulator, SSTW = single-layer sliding time window, SPTW = single-layer parallel time window, DPTW = double-layer parallel time window, ApEn = approximate entropy from steering wheel angle, VCF = velocity change in the car following scene, DCF = distance change from the vehicle in front, SDLP = standard deviation of lane position, PERCLOS = percentage of eyelid closure, SDSWA = standard deviation of steering wheel angle, and PNS = percentage of nonsteering.

decreased dramatically for a new driver because of individual differences [45]. Yan et al. [41] analyzed driver fingerprinting differences in drowsiness-detection models and found that differences in measurement distribution among drivers could be beyond differences caused by drowsiness when all drivers' measurements were mixed to train models, which decreased the correlation between measurements and drowsiness. Therefore, the accuracies of drowsiness-detection models without driver fingerprinting differences were impaired.

For eliminating the negative influence of driver fingerprinting differences on drowsiness-detection accuracy, scholars attempted to use measurements of individual drivers to establish driver-specific models and achieve more accurate drowsiness-detection [39, 41, 43, 46, 47]. Chu et al. [47] recorded naturalistic driving data by actual vehicles, and trained drowsiness-detection models by the speed and lane departure value of individual drivers, based on the combined model of neural network and support vector machine. The drowsiness-identification rate of the model for some drivers exceeded 90%, which verified that driver-specific models could improve the accuracy of drowsiness-detection. You, et al. [46] handled differences in eye movement behavior data among drivers, extracted ocular measurements like eyes aspect ratio, and used ocular measurements of individual drivers to establish the driver-specific drowsiness-detection models. The accuracy of the proposed model reached 94.8%, which outperformed generalized drowsiness-detection models. Wang et al. [39] collected multisource driving behavior data by simulating driving experiments and utilized 23 nonintrusive measurements of individual drivers to build the personalized drowsy driving models based on the multilevel logit model. The results verified that the drowsiness-detection accuracy of the personalized model considering driver

fingerprinting differences was higher. Naurois et al. [48] used data of a group of drivers to train the drowsiness-detection model based on artificial neural networks and used adaptive learning to personalize the drowsiness-detection model for a new driver. The model enhanced drowsiness-detection performance by roughly 40%.

The above driver-specific drowsiness-detection models improved the accuracy of drowsiness-detection, which was attributed to adopting individual drivers' specific drowsiness-detection thresholds. However, in the above driver-specific models, measurements of each driver were still computed by UCPs rather than Individual Drivers' Best Calculation Parameters (IDBCPs), which restrained contributions of measurements for detecting drowsiness. Due to differences in driver fingerprinting among drivers, IDBCPs that made measurements more discriminative for drowsiness were various [18, 49]. For some drivers, IDBCPs of measurement are very different from UCPs. And, using UCPs to compute individual driver's measurements restrains the drowsiness-detection performance of measurements, which decreases the accuracy of driver-specific drowsiness-detection models. Similarly, Zhang et al. [18] found that there were considerable differences in IDBCPs and using the UCPs to calculate measurements might weaken the correlation between measurements and drowsiness levels for some drivers. Therefore, the purpose of this paper was to build a novel model to optimize the measurements' calculation parameters of the individual driver. And, measurements calculated by IDBCPs which could validly distinguish drowsiness were chosen to compose individual drivers' optimal measurement set of drowsy driving. This paper contributes to drowsiness measurement calculations and the establishment of drowsiness-detection models considering driver fingerprinting differences.

The rest of this paper is arranged as follows. In Materials and Methodology, we described the experiment details, measurement calculation method, the extraction model of optimal drowsy driving measurements of individual drivers, and the validation method of the drowsiness-detection advantages of measurements calculated by IDBCP. In Results, we presented the driver fingerprinting differences of measurements using UCPs, the results of the model, and the drowsiness-detection advantages of using IDBCPs to calculate nonintrusive measurements. The discussion of the results, the important contributions, and the application prospects of this paper are introduced in Discussion. And, the summary of research work, research limitations, and future work are given in Conclusions.

## 2. Materials and Methodology

### 2.1. Experimental Design

**2.1.1. Participants.** According to the actual proportion of male and female commercial vehicle drivers locally, 35 male drivers and 5 female drivers were recruited as participants and numbered. All participants were professional drivers with rich driving experience and proficient driving skills. Their age ranged from 34 years old to 57 years old (mean = 46.83, SD = 5.62); the driving age ranged from 3 years to 32 years (mean = 16.53, SD = 6.10). Participants did not have any sleep-related disorders and other physical or mental illnesses, and they took no pharmaceuticals within 1 month before the experiments. On the co-pilot, we arranged one coach with 30 years of driving experience as a safety officer to accompany every participant. The safety officer took countermeasures during emergencies and was trained to inquire and record the participants' self-reported drowsiness level using the Karolinska Sleepiness Scale (KSS). The KSS ranges from one to nine, where one and nine, respectively, represent extremely alert and extremely drowsy [50]. The safety officer has a rich long-distance driving experience and strong anti-fatigue abilities. To prevent the safety officer from using KSS due to excessive fatigue, the safety officer should have a good rest when arriving at the service area. Besides, during driving duration, the safety officer kept relatively alert by drinking coffee and refreshing energy beverages and listening to pop music, etc.

**2.1.2. Devices.** The major experimental devices and routes are shown in Figure 1. The experimental vehicle was modified from the actual vehicle, equipped with a variety of sensors. The multisource naturalistic driving behavior data were collected timely. The three-way HD camera (frame rate was 30 Hz) could capture the driver's facial characteristics, driving operation details, and the traffic environment outside the vehicle including road linearity and traffic flow. The HD cameras were Logitech C910. The driver state sensor (sampling frequency was 60 Hz) was W5Y-DSSV3 (AR-ES5430SM) produced by Seeing Machines, which used an image recognition algorithm to monitor the driver's eyes and eyelid movements. The Mobileye C2-270 (sampling frequency was 15 Hz) could detect lane lines and output the

distance between the vehicle center and the lane lines. The inertial navigation system (sampling frequency was more than 20 Hz) could collect acceleration, velocity, and GPS, which was RT2500 developed by Oxford Technical Solutions. The industrial computer could store raw multiple-source experimental data, which was SIMATIC Rack PC 847B of SIEMENS. The steering wheel angle sensor was LWS 5 of BOSCH (sampling frequency was more than 20 Hz), which collected the steering wheel angle. Besides, other materials were also used, such as the KSS, demographic information scale, etc.

**2.1.3. Implementation Process.** Before experiments, the participants were notified of the important information including the experimental process and the possible risk, and each participant signed informed consent. Participants were trained to understand KSS correctly for ensuring the reliability of self-reported drowsiness levels. Then, participants were familiarized with the experimental vehicles and made sure not to take beverages, such as caffeine, that would affect the measure of drowsiness. On the night before the experiment, participants kept a normal schedule and ensured adequate sleep duration. Before driving, participants were required to report their sleep duration the previous night and current KSS.

In the experiment, participants were required to drive the car along the route shown in Figure 1 with their usual driving habits. The experimental road section was G70 (two-way four-lane) Expressway from Wuhan to Xiangyang, China. The road section was flat and relatively straight. Its traffic flow was simple. So the influence of road shape and other vehicles' disturbance on driving behavior could be eliminated such as steering and lane-keeping, which was out of the scope of the current work. The factor of weather, such as snow, rain, and fog, can also affect naturalistic driving behavior [51], which can interfere with the analysis of the effects of drowsiness on driving behaviors. To control irrelevant variables, all experiments were carried out under the condition of well-lit weather to ensure that the temperature, humidity, and other environmental factors were consistent. Besides, each participant started the experiment at 9 Am on weekdays to avoid the influence of traffic volume changes. All participants entered the G70 highway from the Fuhe toll station. After about 2 hours, participants reached the Suizhou service area and took a 1-hour break. Then, participants drove to Xiangyang north toll station, turned around at this point, and returned to the start point. The total driving time was about 6 hours and the total travel was about 600 kilometers in this experiment. During the process of driving, the noise level was controlled to keep a comfortable environment and prevent noise from influencing the physiological and psychological state of participants. Participants were required to drive at a speed limit of 120 km/h. The same safety officer inquired to record 40 participants' self-report KSS every 5 minutes.

After completing the whole experiment, participants were paid for their participation. And, the experimental designs met the guidelines of the local ethics committee and the



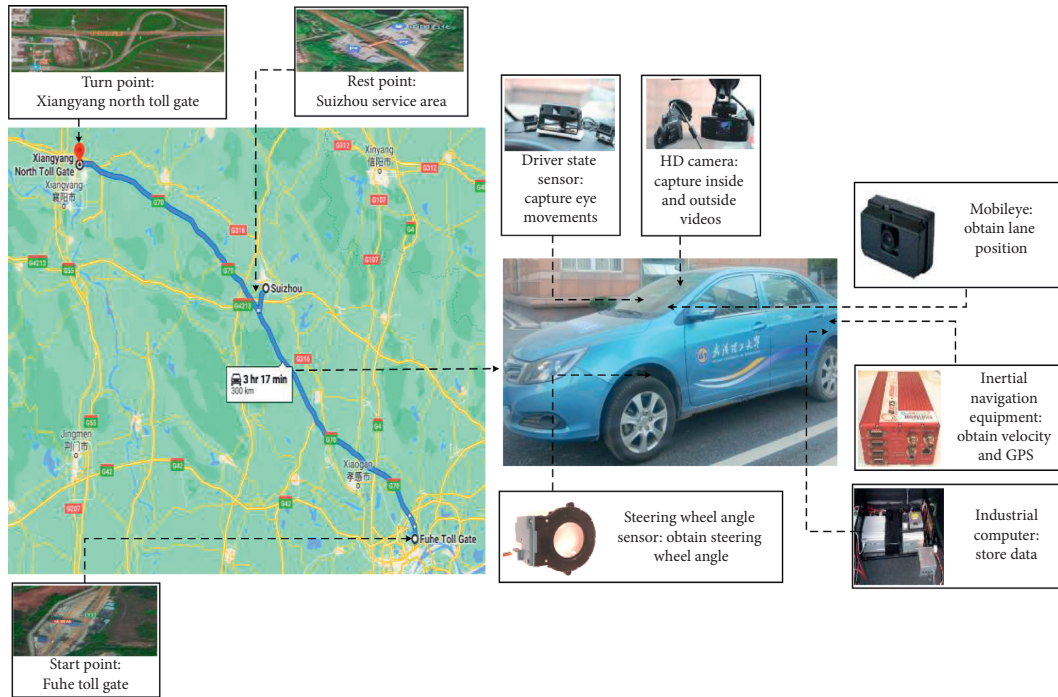


FIGURE 1: Diagram of experiment apparatuses and routes.

personal information of participants was properly secured. Finally, 35 participants' qualified experimental data were obtained. Valid data of 5 participants were not collected due to poor experimental conditions or equipment problems.

**2.2. Data Preprocessing.** According to the flow in Figure 2, raw experimental data were preprocessed to obtain the drowsy driving measurements that meet the requirements of the model.

**2.2.1. Synchronizing Multisource Data.** Based on the timestamp recorded by the sensor, we synchronized multisource data including videos, self-reported KSS, lane position, steering wheel angle, velocity, and so on. Because the driver's drowsiness level was evaluated every 5 minutes, we divided the raw data into samples within 5 minutes to avoid the sample crossing two KSS.

**2.2.2. Choosing Data of the Continuous Driving Scene.** The driving behavior in the continuous driving scene on the motorway was chosen to analyze. The driving operation is more monotonous when the driver is driving continuously on the motorway, and the risks of drowsy driving are greater because of the high speed [39]. According to the videos outside the vehicle, we extracted the driving duration under the continuous driving scene and selected the data of the continuous driving scene.

**2.2.3. Setting Double-Layer Sliding Time Window and Calculating Measurements.** The research in the field experiments found that the drowsy driving state generally lasted

15–75 seconds, while the typical drowsy operation characteristics duration was generally 5–20 seconds [30]. When measurements are calculated by data in the entire drowsy state duration, the calculating average effect covers the drowsy operating characteristics [1, 18]. Thus, the double-layer sliding time window was proposed to divide raw data into samples and calculate measurements.

Firstly, we set the first-layer sliding step ( $S_1$ ) and time window ( $T_1$ ) within every 5 minutes. The experimental data in  $T_1$  were data of a sample, and  $T_1$  was a fundamental unit for calculating measurements and detecting drowsiness. The data in one  $T_1$  could be used to calculate one measurement value, which was one measurement sample. Secondly, within each  $T_1$  time window, we set the second-layer sliding step ( $S_2$ ) and time window ( $T_2$ ). The data in  $T_2$  were used to calculate the measurement. To magnify the drowsy driving characteristics in the sample ( $T_1$ ), the maximum measurement for all  $T_2$  in every  $T_1$  was chosen as the final measurement value of the sample ( $T_1$ ). For instance, when calculating SDLP, if  $T_1 = 60$  s,  $S_2 = 5$  s,  $T_2 = 20$  s, then  $T_1$  included nine  $T_2$  and the lane position data in each  $T_2$  can be used to calculate one SDLP; we chose the maximum SDLP of nine  $T_2$  as the final SDLP of the sample  $T_1$ . The number of  $T_2$  was determined by the size of  $S_2$ . The smaller the  $S_2$  is, the more  $T_2$  can be obtained to select max drowsiness measurement value within  $T_1$ . But if the amount of  $T_2$  was too large, the calculation amount of the model would increase and the calculation speed of the model would reduce. Besides, there are certain gaps between typical fatigue operating features,  $S_2$  cannot be too small. Thus, for reducing the amount of computation, avoiding or missing the best  $S_2$ , and reducing the degree of freedom, we design the optimization range of  $S_2$  to be an integer between 2 and 5.

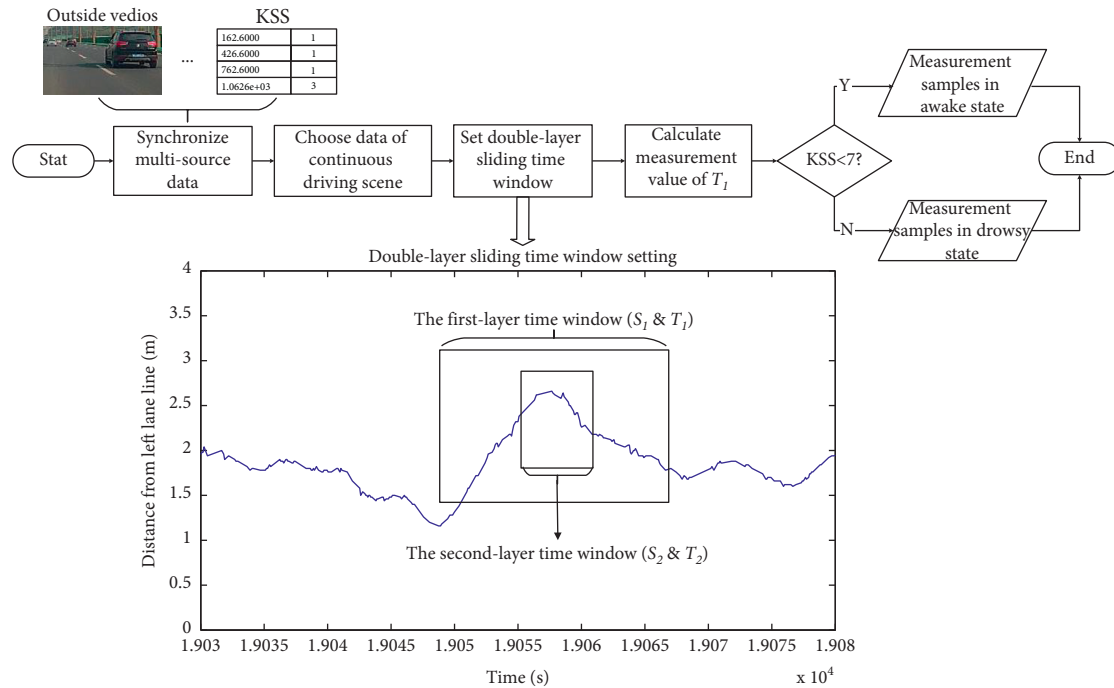


FIGURE 2: The workflow of data preprocessing.

**2.2.4. Dividing Measurement Samples into Samples of Awake State and That of Drowsy State.** The KSS of  $T_1$  was defined as the KSS of measurement samples. Referring to the literature [18, 52], measurement samples with KSS less than 7 belonged to samples of the awake driving state; otherwise, they belonged to samples of the drowsy driving state. In Figure 3, we drew the KSS overtime of 35 participants. The KSS equals zero means stop because of rest or some breakdowns. As shown in Figure 3, with the increase in driving time, KSS gradually increased. Before and after rest in the service area, the KSS peak appeared with the increase of driving time. And after the rest in the service area, KSS decreased to a certain extent, but KSS rapidly increased with the increase of driving time.

**2.3. Measurement Calculation.** Drowsiness can affect the driver's steering wheel rotation, eye movement, and other aspects [12]. According to related literature [12, 18, 39], eight drowsiness-related nonintrusive measurements of different categories were chosen to analyze. In the previous studies, measurements of all drivers were calculated by UCPs. For comparing with the drowsiness-detection contributions of measurements calculated by IDBCPs, we also used suitable UCPs to calculate measurements.

We utilized the Wilcoxon test (introduced in Section 2.4.2) to analyze the differences between measurements of the awake state and that of drowsy state. For a participant, if  $P$  value  $< 0.05$ , the measurement can distinguish drowsy driving [53]. When measurements calculated by UCPs were used to build drowsiness-detection models without considering individual differences, the number of participants whose drowsiness-driving can be detected by the model using unified parameters should be as

many as possible. Given UCPs, we calculated measurements by UCPs and did the Wilcoxon test on individual driver's measurements. The more the participants with  $P$  value  $< 0.05$  are, the better the UCPs are. And for a participant, the same original experimental data were used to calculate UCPs and IDBCPs, so the experimental setting had no influence on the results. Thus, through iterating various UCPs and comparing the number of drivers whose drowsiness can be detected by measurements calculated by different UCPs, we found the suitable UCPs of each measurement which can make the number of participants whose  $P$  value  $< 0.05$  maximum.

According to reference [18, 30], we designed the value ranges of UCPs.  $T_1$  corresponds to the duration of the drowsy state (15 s–75 s), and  $T_2$  corresponds to the duration of the typical characteristics of drowsiness (5 s–20 s). Therefore, the value range of  $T_1$  and  $T_2$  is designed to be 15–75 and 5–20, respectively. The frequencies of sensors were above 20 Hz, and every second contains a lot of data, reflecting the driver's operation details. Therefore, the drowsiness characteristics can be obtained by using the related smaller  $T_2$  to calculate measurements. And, the larger  $T_2$  can contain the nondrowsy data and weaken the drowsiness characteristics.  $S_1$  and  $S_2$  affect the quantity of  $T_1$  and  $T_2$ , respectively. Using small  $S_2$  is convenient for selecting max drowsiness measurement value within  $T_1$ . However, if  $S_2$  is too small, it will increase the number of  $T_2$  and increase the amount of model computation. Also, there are gaps among typical fatigue operating features. To make sure the best quality and reduce the calculation amount, we set the range of  $S_1$  and  $S_2$  to be 2–5 and 7–10, respectively, to balance the amount of computation and reduce the degree of freedom. Table 2 is about the important information of nonintrusive measurements.

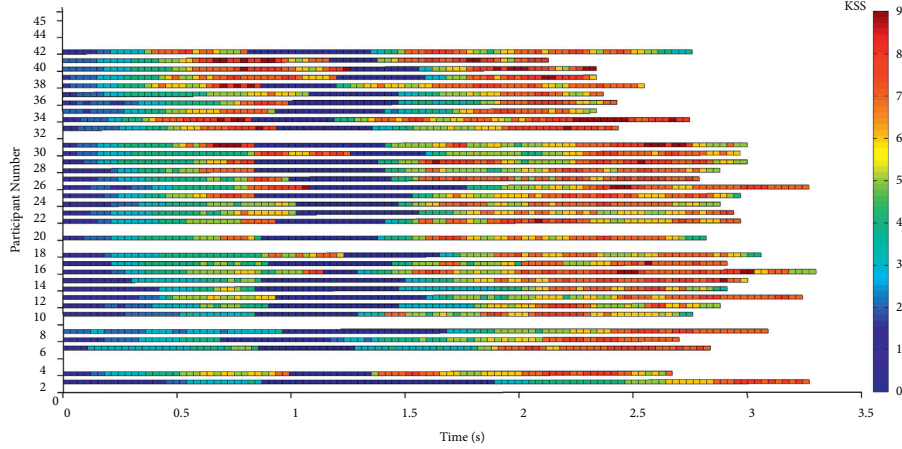


FIGURE 3: Graph of KSS changes of all participants over time.

TABLE 2: The information on nonintrusive drowsiness-related measurements.

Data source	Order	Symbol (unit)	UCPs (unit)
Mobileye	①	SDLP (m)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 18$ (s)
Steering wheel angle sensor	②	SRR (/)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 18$ (s), Thr = 6 (degree)
	③	SDSWM (degree/s)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 2$ (s), $T_2 = 15$ (s)
	④	VSA (/)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 18$ (s)
Inertial navigation sensor	⑤	SDLA ( $m^2/s$ )	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 14$ (s)
	⑥	SDTV (m/s)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 15$ (s)
Driver state sensor	⑦	MPE (/)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 3$ (s), $T_2 = 15$ (s)
	⑧	SDPE (/)	$S_1 = 6$ (s), $T_1 = 45$ (s), $S_2 = 4$ (s), $T_2 = 18$ (s)

SRR = steering reversal rate, Thr = reversal threshold, / = dimensionless measurements have no units, SDSWM = standard deviation of steering wheel movement, VSA = variation of steering wheel angle, SDLA = standard deviation of longitudinal acceleration, SDTV = standard deviation of transverse velocity, MPE = mean of PERCLOS, and SDPE = standard deviation of PERCLOS.

Calculation formulas of some complicated measurements are as follows.

② SRR measures the frequency of sampling points whose angle exceeds the wheel angle threshold, which can reflect the stability of steering wheel control [18].

$$SRR = \frac{N_{thr}}{N_{all}} \quad (1)$$

$N_{thr}$  is the number of sampling points whose angle exceeds the thr and  $N_{all}$  is the total number of sampling points in the time window.

④ VSA can eliminate the influence of road curvature on steering wheel angle to some extent, which reflects the relationship between drowsiness and steering wheel angle [54].

$$VSA = \frac{SDSWA}{MSWA} \quad (2)$$

MSWA and SDSWA are, respectively, the mean value and standard deviation of the steering wheel angle in the time window.

⑧ SDPE describes the variation of PERCLOS in the time window, which reflects the blink frequency.

$$SDPE = \sqrt{\sum_{i=1}^{n_{pe}} \left( PE_i - \frac{1}{n_{pe}} \sum_{i=1}^{n_{pe}} PE_i \right)^2 \frac{1}{n_{pe} - 1}} \quad (3)$$

$PE_i$  is the PERCLOS at sampling point  $i$ , and  $n_{pe}$  is the total number of PERCLOS sampling points in the time window. The PERCLOS is directly output by DSS. Within a certain period, DSS calculates the proportion of time the eyes are at least 80 percent closed. Then, for increasing the volume of data, DSS obtains the PERCLOS of each sampling point through interpolation. And, the sampling frequency of PERCLOS is about 60 Hz. Therefore, we can calculate the mean and standard deviation of PERCLOS in  $T_2$ .

#### 2.4. Extraction Model of Individual Driver's Optimal Drowsiness-Detection Measurements

2.4.1. *Structure of the Model.* The framework for the extraction model of optimal drowsiness-detection measurements of individual drivers based on the Wilcoxon test and genetic algorithm (GA) is shown in Figure 4.

As shown in Figure 4, the model is divided into three parts. Measurements of the individual driver are input into this model to obtain IDBCPs and compose the individual driver's optimal drowsiness-detection measurement set. In part A, through performing the Wilcoxon test on measurements of the awake state and that of drowsy states, we obtain  $|Z\text{-statistics}|$  ( $|Z|$ ) representing the drowsiness-detection performance of measurements and construct the

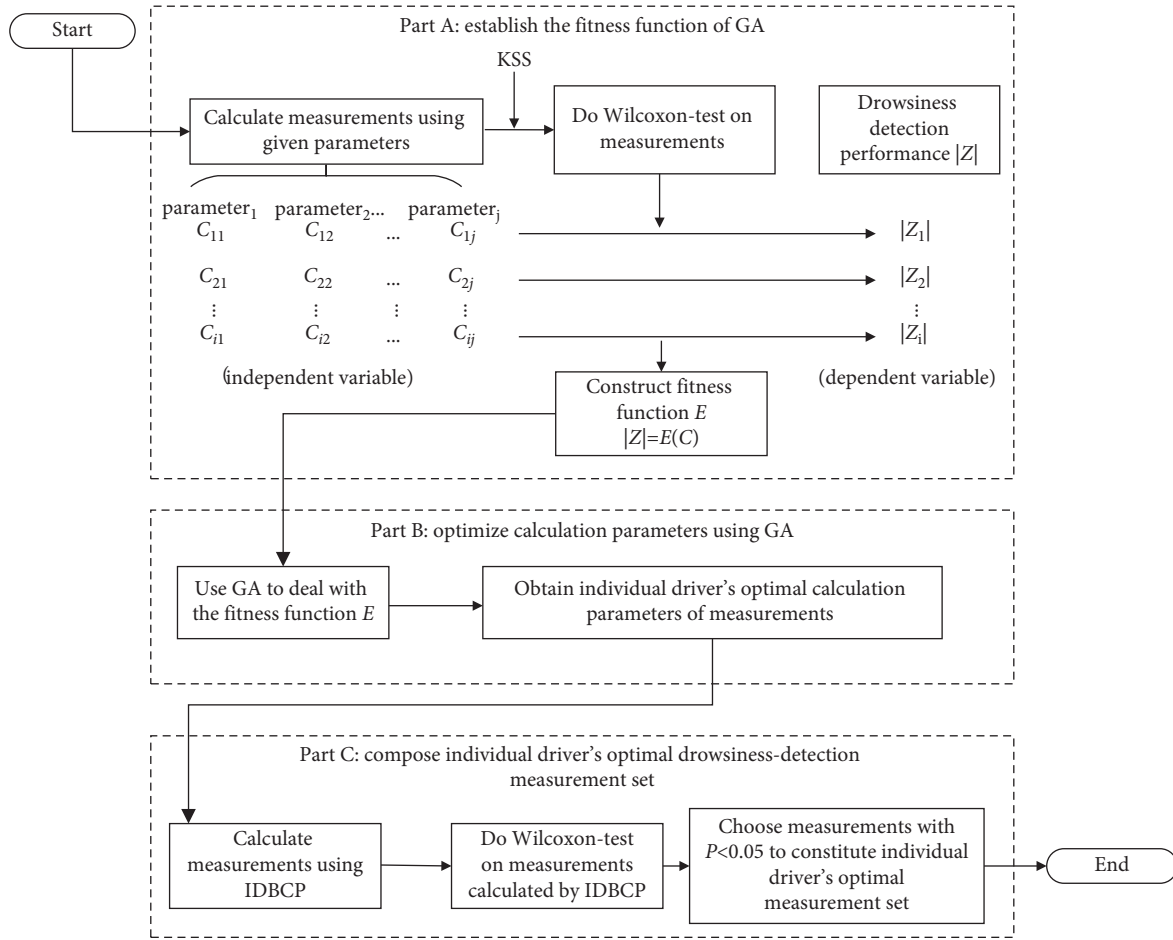


FIGURE 4: Extraction model of optimal drowsy driving measurements of individual drivers.

fitness function reflecting the correspondence between the calculation parameters and measurements' drowsiness-detection performance. In part B, the GA is used to optimize the fitness function and obtain the measurements' IDBCPs. In part C, we complete the Wilcoxon test on measurements that are computed by IDBCPs. We choose the measurements whose  $P$  value  $< 0.05$  to compose the individual driver's optimal drowsiness-detection measurement set.

**2.4.2. Establishing Fitness Function of GA.** Firstly, the measurement samples are computed by the given calculation parameters. And, according to KSS, measurement samples are divided into samples of the awake state or that of the drowsy state.

Secondly, we perform the Wilcoxon test on measurements of the awake state and that of the drowsy state and obtain  $Z$ -statistics. Measurement sample sizes of the awake state and that of the drowsy state are different and samples are not normally distributed. Therefore, the Wilcoxon test was used to analyze the differences between the measurement samples of the awake state and that of the drowsy state. The Wilcoxon test is to analyze whether differences between unpaired samples from two groups are statistically significant without requesting normal distribution [53]. The bigger

the  $|Z|$ , the greater the difference between measurement samples of the awake state and drowsy state. Thus,  $|Z|$  represents the drowsiness-detection performance of measurements. The method of calculating  $Z$ -statistics is as follows.

Measurement samples in the awake driving state are  $A: A_1, \dots, A_i, \dots, A_n$ , and measurement samples in the drowsy driving state are  $D: D_1, \dots, D_j, \dots, D_m$ . The sample size of the awake state and that of the drowsy state is, respectively,  $n$  and  $m$ . Measurement samples of awake state and that of drowsy state are mixed. The mixed measurement samples are sorted in the ascending order.

$$\begin{aligned}
 W_A &= \sum_{i=1}^n R_i W_D = \sum_{j=1}^m R_j W_{DA} = W_A - \frac{m(m+1)}{2} W_{AD} \\
 &= W_D - \frac{n(n+1)}{2}, \\
 U &= \min\{W_{DA}, W_{AD}\}, \\
 Z &= \frac{U - mn/2}{\sqrt{mn(n+m+1)/12}} \sim N(0, 1).
 \end{aligned}$$

(4)



$R_i$  and  $R_j$  are the rank in mixed samples of the awake sample  $A_i$  and the drowsy sample  $D_j$ , respectively.  $N(0, 1)$  means that the data are normally distributed.

Finally, the fitness function is established as follows:

$$|Z| = E(C). \quad (5)$$

$C(C_1, C_2, \dots, C_j)$  is a vector consisting of independent variables and  $C_j$  was the  $j$ -th calculation parameter of the measurement. Take SRR in Table 2 as an example; the calculation parameters of SRR are  $S_1, T_1, S_2, T_2$ , and Thr. Therefore,  $C(C_1, C_2, \dots, C_j)$  of SRR is  $C(S_1, T_1, S_2, T_2, \text{Thr})$ . We can get the corresponding  $|Z_i|$  when a set of calculation parameters  $C_i(C_{i1}, C_{i2}, \dots, C_{ij})$  is given. Therefore, the fitness function reflects the relationship between the calculation parameters and the drowsiness-detection performance ( $|Z|$ ) of the measurement. And, the larger the fitness value ( $|Z|$ ) is, the better the calculation parameters of the measurement are.

**2.4.3. Optimizing Calculation Parameters Using GA.** The genetic algorithm is a method of searching for the optimal solution based on population intelligence with advantages of simplicity, high efficiency, and good robustness [55]. Thus, we used GA to optimize the fitness function ( $|Z| = E(C)$ ) to obtain measurements' IDBCPs. Chromosomes are individuals, and multiple individuals constitute a population. The flow chart of GA is shown in Figure 5. And the main steps of GA are introduced as follows:

Encoding measurement's calculation parameters: a chromosome consists of calculation parameters of measurements that need to be optimized. We use binary coding and decoding because of its advantages of high search efficiency and easy convergence, and convert calculation parameters into binary strings of different lengths according to the solution space and accuracy of the calculation parameters. For instance, when we take SDLP in Table 2, the calculation parameters of SDLP are  $S_1, T_1, S_2, T_2$ , and for SDLP, chromosome =  $[S_1, T_1, S_2, T_2] = (6, 45, 4, 18)_{10} = (110, 101101, 100, 10010)_2$ .

Calculating fitness value: the chromosome which represents measurement's calculation parameters is input into the function ( $|Z| = E(C)$ ). Through the Wilcoxon test on measurements, the fitness value ( $|Z|$ ) of the chromosome is obtained.

Choosing reproduction: chromosomes with larger fitness values are more likely to be selected for reproduction (crossover or mutation) [55]. The probability of the chromosome being selected equals to  $|Z_i| / \sum |Z_i|$ .  $Z_i$  is the fitness value of chromosome  $i$ . And, we obtain better chromosomes by roulette selection.

Crossover: two chromosomes are randomly selected to crossover according to the cross probability ( $P_{cr}$ ), and then a pair of nodes are randomly selected for the exchange to generate two new chromosomes.

Mutation: a chromosome is randomly select to mutate according to the mutation probability ( $P_{mu}$ ). And, then a node on its binary string is randomly selected to replace it with an allele value.

Relative process parameters of GA: learn from related research [55], parameters of GA are designed as follows,  $N = 50$  ( $N$  is the population size),  $P_{cr} = 0.6$ ,  $P_{mu} = 0.01$ .

Termination condition: after the genetic algebra reaches 200, GA is ended. Finally, by decoding, the binary optimal solution is converted to decimal, and IDBCPs of measurements are output.

In this paper, for fusing multiple measurements of a sample ( $T_1$ ) to identify drowsiness, all measurements should have the same  $S_1$  and  $T_1$ . First, some groups ( $S_1, T_1$ ) are generated by enumeration. Then, all measurements use the same ( $S_1, T_1$ ), and the GA is used to optimize the remaining calculation parameters to obtain the maximum  $|Z_{max}|$  of each measurement under the same ( $S_1, T_1$ ). Finally, ( $S_1, T_1$ ) with the largest sum of all measurements'  $|Z_{max}|$  is chosen as the best ( $S_1, T_1$ ) of all measurements. Finally, ( $S_1, T_1$ ) with the largest sum of all measurements'  $|Z_{max}|$  is chosen as the best ( $S_1, T_1$ ) of all measurements, which together with other optimal parameters constitutes the measurements' IDBCPs.

**2.4.4. Composing Individual Driver's Optimal Drowsiness-Detection Measurement Set.** Although we optimize the calculation parameters of individual driver's measurements in part B, whether the measurements calculated by IDBCPs can effectively detect drowsy driving needs further test. Thus, we perform Wilcoxon test on measurements calculated by IDBCPs, and then measurements with  $P$ -value  $< 0.05$  are selected to compose the optimal drowsiness-detection measurement set of the individual driver.

**2.5. Validation Model of Individual Driver's Measurements Calculated by IDBCPs.** For verifying the drowsiness-detection advantages of measurements' IDBCPs, for the individual driver, the multiple nonintrusive measurements of individual drivers were integrated to build driver-specific drowsiness-detection models based on the Fisher discriminant algorithm. We, respectively, used drowsiness-detection measurements calculated by UCPs and IDBCPs to build two kinds of driver-specific drowsiness-detection models named DF\_U and DF\_I. The drowsiness-detection advantages of using IDBCPs are illustrated by comparing drowsiness-detection accuracies of DF\_U and DF\_I.

Fisher discriminant algorithm is an effective classification method, which simplifies the problem by projecting the points in the high-dimensional space [56].

$$\begin{bmatrix} DS_1 \\ DS_2 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} C^T + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (6)$$

$$P_i = \frac{e^{DS_i}}{e^{DS_1} + e^{DS_2}}, \quad i = 1, 2.$$

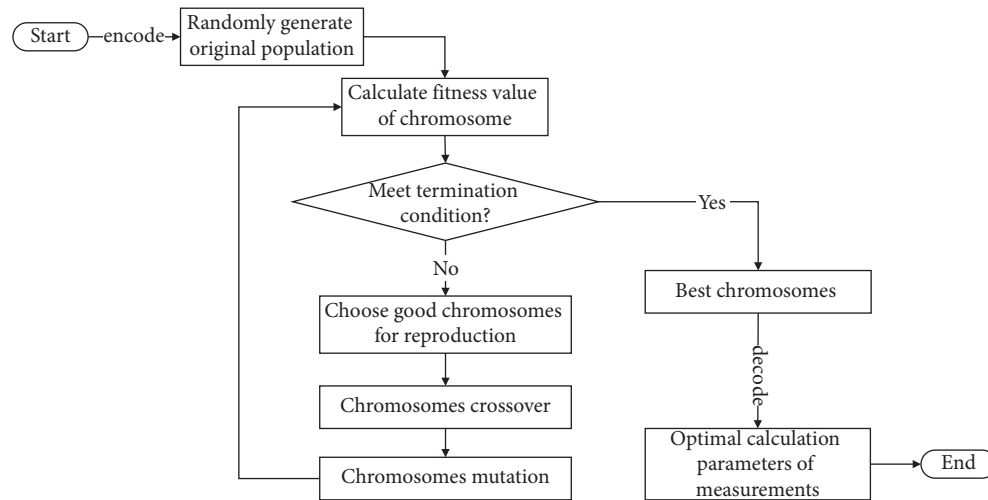


FIGURE 5: The flow chart of GA.

$DS_1$  and  $DS_2$  are the function values of the awake state and drowsy state, respectively.  $\alpha(\alpha_1, \alpha_2, \dots, \alpha_6), \beta(\beta_1, \beta_2, \dots, \beta_6), b_1, b_2$  are the algorithm coefficient, which is obtained by using training samples to train the model.  $P_i (i = 1, 2)$  is the drowsiness-detection measurement of individual drivers such as SDLP and SRR.  $P_i (i = 1, 2)$  is the probability that the sample belongs to the awake state or drowsy state. The driving state with a higher probability is the identification result of the model.

### 3. Results

To save paper space, 6 participants (No. 5, No. 11, No. 14, No. 18, No. 23, and No. 29) were randomly selected as examples to display the related results in this section.

**3.1. Driver Fingerprinting Differences of Measurements Using UCPs.** We took SDLP as an example to display driver fingerprinting differences of measurements calculated by UCPs. Figure 6 shows the radar charts composed of the quartiles of the SDLP of 6 participants in the awake state and drowsy state.

In Figure 6, all closed lines were not circular but irregular polygons. It meant that the distributions of SDLP calculated by the same calculation parameters of 6 participants were different, which was the outcome of driver fingerprinting differences in lane control behavior. There were differences in the median of SDLP among participants in the awake state and drowsy state. For instance, in the awake state, the medians of No. 11 and No. 18 were maximum (0.35) and minimum (0.20), respectively. And, in the drowsy state, the medians of No. 11 and No. 18 were maximum (0.41) and minimum (0.24), respectively. Besides, the ranges of SDLP among participants were also different. In the awake state, the differences between the upper and lower quartiles of No. 11 and No. 18 were the maximum (0.18) and minimum (0.07), respectively. In the drowsy state, the differences between the upper and lower quartiles of No. 14 and No. 18 were the maximum (0.21) and minimum (0.12), respectively.

**3.2. Optimal Drowsiness-Detection Measurements of Individual Participants.** All samples were divided into the training sample set and testing sample set, with the proportions of 70% and 30%, respectively. Using stratified random sampling, 70% of the data in awake and drowsy states were, respectively, selected to form the training sample set, and the remaining data were the testing sample set. The stratified random sampling was repeated three times, and the final result was taken as the average value of repeated calculations. Table 3 presents the results of models in Figure 4 of the example participants. And in Table 3,  $|Z|$  is  $|Z\text{-statistics}|$  of measurements calculated by IDBCPs, and \* means  $P$  value  $< 0.05$ .

As shown in Table 3, most measurements using IDBCPs could distinguish drowsy driving ( $P$  value  $< 0.05$ ). However, for some participants, although the measurements' calculation parameters were optimized, the measurements calculated by IDBCPs could not still distinguish drowsy driving ( $P$  value  $< 0.05$ ). For example, the  $P$  value of SDSWM of No. 11 exceeded 0.05. Therefore, SDSWM could not be selected to compose an optimal drowsiness-detection measurement set of No. 11.

Participant No. 14 was chosen as an example to display the optimization process of SRR's calculation parameters. The training samples of SRR were input into the model in Figure 4. As is shown in Figure 7, the fitness value gradually increased with iterations, which indicated that the model could optimize the calculation parameters and improve the drowsiness-identification performance of measurements.

**3.3. Comparison of Measurements Calculated by IDBCPs and UCPs.** The testing samples were used to analyze the change of drowsiness-detection performance of measurements calculated by IDBCPs. In Table 4, we took SDSWM as an example, and  $|Z|$  of SDSWM calculated by UCPs ( $|Z|\text{-UCPs}$ ) and IDBCPs ( $|Z|\text{-IDBCPs}$ ) were, respectively, listed.

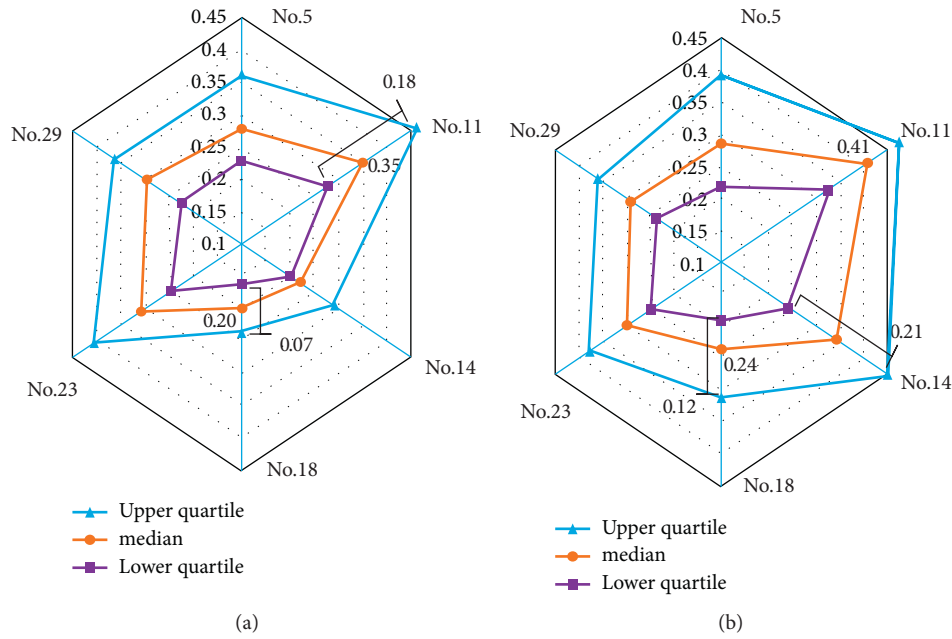


FIGURE 6: The SDLP calculated by the UCPs radar chart of 6 participants. (a) SDLP of the awake driving state. (b) SDLP of the drowsy driving state.

TABLE 3: The IDBCPs and  $|Z|$  of measurements for example participants.

Participant number		No. 5 (17)	No. 11 (23)	No. 14 (31)	No. 18 (30)	No. 23 (35)	No. 29 (41)
Measurements		$S_1 = 7, T_1 = 62$	$S_1 = 7, T_1 = 66$	$S_1 = 7, T_1 = 54$	$S_1 = 7, T_1 = 56$	$S_1 = 8, T_1 = 65$	$S_1 = 7, T_1 = 50$
SDLP	$S_2$	3	4	5	2	3	5
	$T_2$	20	15	20	18	14	17
	$ Z $	4.65*	4.44*	6.84*	6.54*	2.83*	1.17
SRR	$S_2$	5	2	3	3	2	2
	$T_2$	20	20	20	18	15	16
	Thr	5	6	6	7	6	7
	$ Z $	4.34*	2.34*	3.23*	2.64*	4.07*	6.03*
SDSWM	$S_2$	3	5	5	2	2	2
	$T_2$	20	14	19	19	14	20
	$ Z $	7.05*	1.64	7.51*	6.40*	9.30*	7.65*
VSA	$S_2$	3	5	3	5	5	5
	$T_2$	20	18	14	17	20	14
	$ Z $	3.94*	2.95*	9.46*	2.05*	2.05*	3.93*
SDLA	$S_2$	3	3	3	4	3	2
	$T_2$	19	20	14	14	18	20
	$ Z $	5.89*	5.42*	2.57*	3.91*	2.59*	3.51*
SDTV	$S_2$	3	5	3	2	2	2
	$T_2$	20	17	16	14	15	20
	$ Z $	8.67*	1.93	7.54*	10.84*	4.75*	3.82*
MPE	$S_2$	3	2	2	2	4	4
	$T_2$	14	16	16	14	14	19
	$ Z $	6.22*	4.32*	10.79*	5.14*	16.63*	2.83*
SDPE	$S_2$	3	5	4	3	3	4
	$T_2$	20	14	20	20	18	14
	$ Z $	3.92*	3.02*	5.54*	1.87	8.93*	1.73

In Table 4, for every participant,  $|Z|$ -IDBCPs was bigger than  $|Z|$ -UCPs, which meant that using IDBCPs to calculate SDSWM could improve the contribution of SDSWM for

drowsiness detecting. However, the improvement extent for participants was various, that of No. 18 and No. 29 was, respectively, maximum (2.33) and minimum (0.66). It could

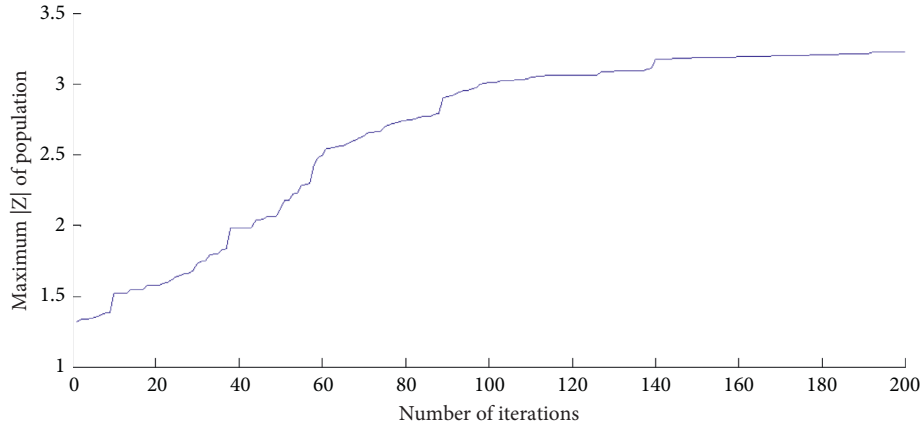


FIGURE 7: The maximum  $|Z|$  of SRR in iterations of GA.

TABLE 4:  $|Z|$  of SDSWM calculated by UCPs and IDBCPs, respectively.

Participant	$ Z $	No. 5	No. 11	No. 14	No. 18	No. 23	No. 29
$ Z _{\text{IDBCPs}}$		4.60	1.41	7.37	6.16	8.16	6.49
$ Z _{\text{UCPs}}$		3.71	0.67	6.59	3.83	6.95	5.84
$\Delta Z $		0.89	0.74	0.78	2.33	1.21	0.66

$$\Delta|Z| = |Z|_{\text{IDBCPs}} - |Z|_{\text{UCPs}}.$$

be seen that the necessities of using IDBCPs for different participants were different, and it was more urgent for No. 18 to use IDBCPs to calculate SDSWM.

To study the distribution differences between measurements calculated by UCPs and that calculated by IDBCPs, in Figure 8, we chose No. 18 to, respectively, draw the boxplots of SDSWM calculated by UCPs and IDBCPs in the awake state and drowsy state.

In Figure 8, whether the SDSWM was calculated by UCPs or IDBCPs, the median of the SDSWM in the drowsy state was greater than that in the awake state. Although there is still an overlap between SDSWM of awake state and that of drowsy state, by performing the Wilcoxon test on SDSWM of awake state and that of drowsy state,  $P$  value was less than 0.05, which illustrated differences between SDSWM in waking state and SDSWM in the drowsy state was statistically significant, and drowsy driving could be distinguished by using SDSWM. It indicated that the drowsiness impaired the ability to control the steering wheel and reduced the stability of the steering wheel movement, which was consistent with previous study conclusions [12]. Besides comparing with SDSWM calculated by UCPs, the overlap decreases when using IDBCPs to calculate SDSWM. The difference between the median of SDSWM calculated by UCPs of the awake state and that of the drowsy state was 0.05, while the difference between the median of SDSWM calculated by IDBCPs of the awake state and that of the drowsy state was 0.07. The result showed that when measurements were calculated by IDBCPs, distribution differences between measurements in the awake state and that in the drowsy state were more significant, and its drowsiness-detection performance was stronger.

**3.4. Verification of Individual Driver's Optimal Drowsiness-Detection Measurements.** Drowsiness-detection advantages of measurements calculated by IDBCPs were verified by comparing drowsiness-detection accuracies of DF\_U and DF\_I. For the individual participant, the optimal drowsiness-detection measurements calculated by IDBCPs were fused to build DF\_I, and the same drowsiness-detection measurements calculated by UCPs were also fused to build DF\_U. For individual drivers, according to stratified sampling, 80% of the samples were, respectively, extracted from measurements in the awake state and that in the drowsy state to compose the training sample set, and the rest constituted the testing sample set. The training samples were used to train driver-specific drowsiness-detection models based on the Fisher discriminant algorithm. In Table 5, for DF\_U and DF\_I, we displayed the mean, standard deviation, minimum, and maximum of drowsiness-detection results across all participants.

As shown in Table 5, the related drowsiness-detection results of DF\_U were better than that of DF\_I. The mean drowsiness-detection accuracy across 35 participants of DF\_I was 91.06%, which was higher than that of DF\_U (85.25%). The results indicated building driver-specific models based on measurements calculated by IDBCPs could improve drowsiness-detection accuracies.

For every participant, Figure 9 shows the comparison of the drowsiness-detection accuracy of DF\_U and DF\_I. In Figure 9, for each example participant, the accuracy of DF\_I was higher than that of DF\_U. Besides, for different participants, the accuracy improvements equalling the accuracy of DF\_I minus the accuracy of DF\_U were various. The accuracy improvement of No. 6 was maximum with the value of 8.11% (DF\_U = 81.73%, DF\_I = 89.84%), whereas that of No. 34 was minimum with the value of 1.56% (DF\_U = 86.89%, DF\_I = 88.45%). The possible interpretation was that the IDBCPs of some participant were similar to UCPs and the improvement of the drowsiness-detection performance of measurements using IDBCPs is limited.



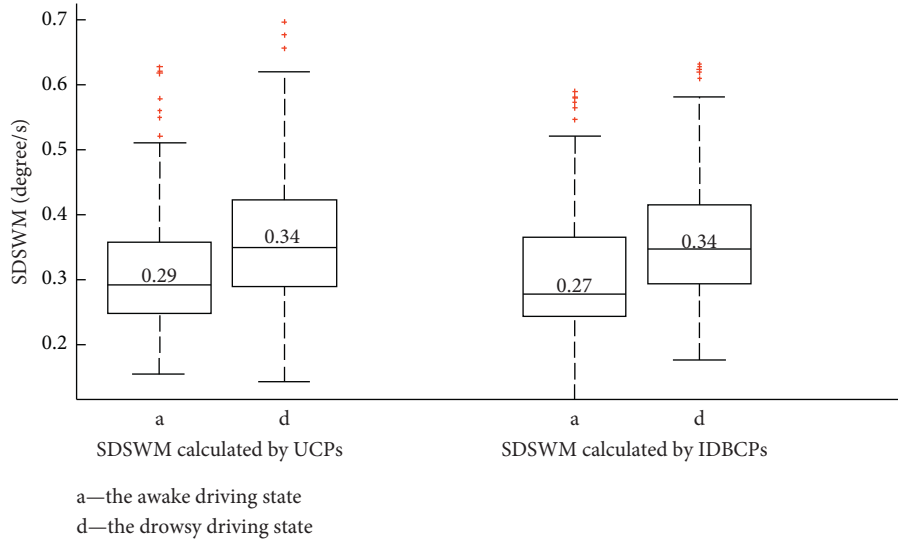


FIGURE 8: The boxplots of SDSWM calculated by UCPs and IDBCPs.

TABLE 5: Results of drowsiness-detection across all participants.

		Mean (%)	Standard deviation (%)	Minimum (%)	Maximum (%)
DF_U	Accuracy	85.25	2.35	79.4	90.2
	Sensitivity	87.5	2.43	81.33	91.46
	Specificity	84.15	2.54	78.32	89.6
	F1	79.56	3.1	71.76	85.71
DF_I	Accuracy	91.06	2.93	81.63	95.97
	Sensitivity	93.39	3.06	83.95	97.78
	Specificity	89.92	3.02	80.49	95.6
	F1	87.38	3.91	75.14	94.51

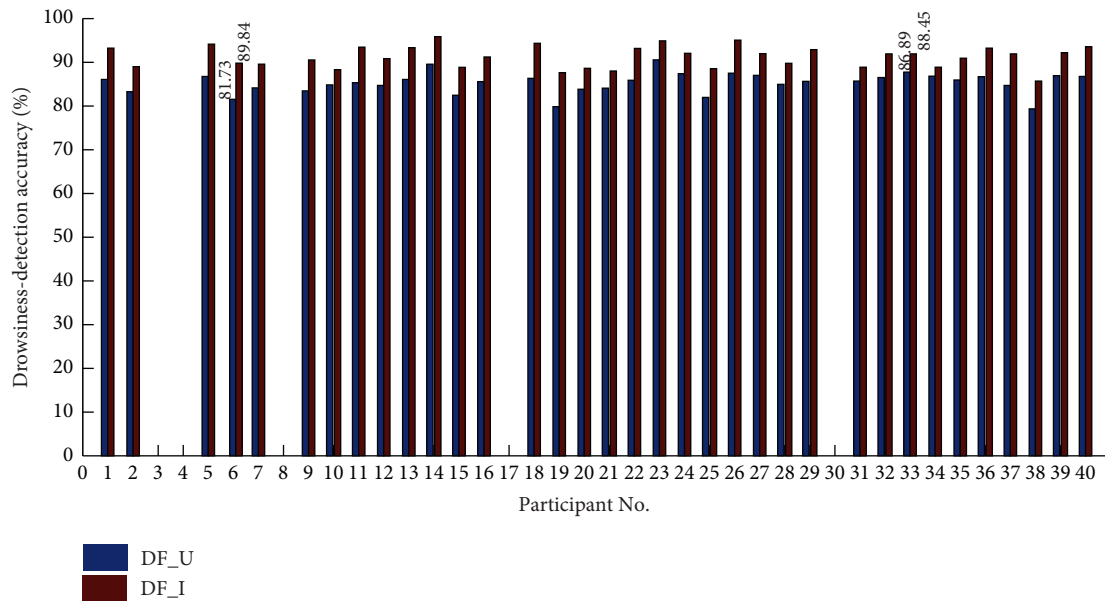


FIGURE 9: The drowsiness-detection accuracy of DF\_U and DF\_I.

## 4. Discussion

By summarizing the above results, important insights and contributions to the development of drowsiness-detection methods considering driver fingerprinting differences were obtained. There are individual differences in data distribution of measurements using UCPs (Figure 6), which is consistent with the previous research [37]. The reason for this phenomenon is that there are differences in drivers' raw driving behavior data, which is the typical manifestation of driver fingerprinting. Similar to the studies [33, 34], it is indicated that individual drivers have unique characteristics in driving behavior and physiology, and there are differences in the driver fingerprints among the drivers. Drowsiness-detection thresholds are the average of all drivers when all drivers' mixed measurement data are used as a whole to train models. And, driver fingerprinting differences make drowsiness-detection of the individual driver to suffer from systematic error. For example, the individual driver's drowsiness-detection threshold is below the average threshold (like No. 18 in Figure 6), the driver's drowsiness cannot be detected when measurements exceed this individual driver's drowsiness-detection threshold but not reach the average threshold of all drivers. Therefore, it is worthy to study accurate drowsiness-detection methods considering driver fingerprinting differences.

The major achievement of this paper is to propose a model to optimize individual driver's calculation parameters of measurements considering driver fingerprinting differences and extract individual driver's optimal drowsiness-detection measurements. It is found that the IDBCPs of measurements among drivers are significantly different (Table 3), which is attributed to driver fingerprinting differences among drivers. Figure 8 confirms that using IDBCPs to calculate individual drivers' measurements can enhance the contribution of measurements for detecting drowsy driving. Similar to existing studies [18, 32, 36], every driver has unique driver fingerprinting characteristics, and the effect of drowsiness on the driving behavior of each driver is also different [10], which causes many individual differences in the raw data related to the drowsiness. For example, the time length of the drowsy state, duration of typical drowsy driving behavior characteristics, and distribution of drowsiness-related driving behavior data of each driver are different. Without considering the driver fingerprinting differences, using UCPs to calculate measurements causes measurements to fail to accurately reflect drowsy driving behavior characteristics of individual drivers, which limits the drowsiness-detection performance of nonintrusive measurements at the individual driver level. Thus, it is suggested to optimize measurements' calculation parameters of individual drivers to obtain IDBCPs suitable for individual driver's driver fingerprinting characteristics at different drowsiness levels, which can strengthen the sensitivity of measurements to drowsiness.

Comparison results of drowsiness-detection accuracy in Figure 9 illustrate that DF\_I outperforms DF\_U, which verifies the drowsiness-detection advantages of using IDBCPs to calculate individual driver's measurements. The

reason is that DF\_I utilizes measurements calculated by IDBCPs with stronger drowsiness-identification performance to build models. The previous study establishes the driver-specific drowsiness-detection model in a simulated driving environment using measurements calculated by UCPs [39]. Although the accuracy of DF\_U (84.16%) is lower than that of previous research (88.60%), there are many interferences in the real driving environment increasing complexities and difficulties of detecting drowsiness. Thus, the drowsiness-detection accuracy of DF\_U is ideal. Moreover, the average drowsiness-detection accuracy of DF\_I is higher than that of the previous research [39], and accuracies of some drivers reach more than 93.00% (No. 14 and No. 23 in Figure 9). It is demonstrated that the drowsiness of individual drivers can be reliably detected with higher accuracy using nonintrusive measurements calculated by IDBCPs, which is also inferred in the previous study [18].

As mentioned in the existing studies, the calculation and selection of measurements is an important factor affecting drowsiness-detection accuracy [26, 39]. Therefore, the model and results of this paper can be applied to develop personalized anti-drowsiness systems in real-world conditions. Individual driver's optimal drowsiness-detection measurements calculated by IDBCPs should be used to train driver-specific drowsiness-detection models. According to the results in Figure 9, the drowsiness-detection accuracy can be improved by using measurements calculated by IDBCPs having higher drowsiness-detection performance. It is worth highlighting that for some drivers whose accuracy improvement of DF\_I is high, it is more indispensable to build the driver-specific drowsiness-detection models using measurements calculated by IDBCPs.

## 5. Conclusions

Due to driver fingerprinting differences, the IDBCPs matching individual driver's behavior characteristics may be very different from UCPs, and using IDBCPs to calculate measurements enhances measurements' drowsiness-detection performance for individual drivers. Therefore, the purpose of this paper was to propose a model to optimize measurements' calculation parameters of individual drivers. Firstly, the naturalistic driving data and KSS of 35 participants were collected through field experiments. Eight nonintrusive measurements related to drowsiness were chosen and computed using the double-layer slide time window. Then, based on Wilcoxon test and GA, we established a model (Figure 4) to extract individual drivers' optimal drowsiness-detection measurements. Finally, based on the Fisher discriminant algorithm, the two kinds of driver-specific drowsiness-detection models were built using measurements calculated by UCPs and IDBCPs, respectively. And, drowsiness-detection advantages of using IDBCPs to calculate measurements were illustrated.

In the present paper, we obtained individual driver's optimal drowsiness-detection measurements calculated by IDBCPs, and the results of example participants are shown

in Table 3. Figure 6 verifies that there are individual differences in nonintrusive measurements calculated by UCPs, which is attributed to the driver fingerprinting differences among drivers. This paper verifies that for individual drivers, the drowsiness-detection performance of measurements calculated by IDBCPs is better than that of measurements calculated by UCPs. In addition, the results showed that the average drowsiness-detection accuracy of DF\_I was 6.25% higher than that of DF\_U, which indicated that using nonintrusive measurements calculated by IDBCPs to establish the driver-specific drowsiness-detection models improved drowsiness-detection accuracy at the individual driver level.

Admittedly, there are some limitations in this paper. For instance, we used participants' self-report KSS as ground truth for drowsiness, and the drowsy levels were relatively subjective. Besides, the amount of experimental data of some individual drivers was not very sufficient which led to insufficient model training and not very high drowsiness-detection accuracy for some participants. For a new user, firstly, we need to label the data with drowsiness and awakeness, which is cumbersome and a limitation in practice. However, for some special drivers such as dangerous goods transport drivers and long-distance bus drivers, it is still very worthwhile to establish a personalized model based on measurements calculated by IDBCPs to improve the drowsiness-detection accuracy.

In the future, we will adopt the more objective and precise standard as the ground truth for drowsiness such as electroencephalogram. And, higher performance models such as artificial neural networks and Bayesian networks will be used to build drowsiness-detection models to reach better accuracies. Ultimately, this research can provide references for the calculation of drowsiness measurements using naturalistic driving data, and guide the establishment of drowsiness-detection models considering driver fingerprinting differences, which can accelerate the development of personalized anti-drowsiness driving active safety systems in vehicles. Meanwhile, the efforts highlight the advantages of studying driver fingerprinting differences and promote the application of driver fingerprinting in the field of dangerous driving behavior.

### Data Availability

The raw data used in this study are available from the corresponding author on request.

### Conflicts of Interest

The data indicated in the findings have not been made available due to data privacy.

### Acknowledgments

This research was supported by the National Key Research Program of China (2019YFB1600800); National Natural Science Foundation of China (52072289 and U1764262).

### References

- [1] G. Adamos and E. Nathanail, "Testing the effectiveness of objective and subjective predictors of driving behavior under fatigue," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 8, pp. 343–352, 2019.
- [2] G. Zhang, K. K. W. Yau, X. Zhang, and Y. Li, "Traffic accidents involving fatigue driving and their extent of casualties," *Accident Analysis & Prevention*, vol. 87, pp. 34–42, 2016.
- [3] Akerstedt, "Consensus statement: fatigue and accidents in transport operations," *Journal of Sleep Research*, vol. 9, no. 4, p. 395, 2000.
- [4] A. Moradi, S. Nazari, and K. Rahmani, "Sleepiness and the risk of road traffic accidents: a systematic review and meta-analysis of previous studies," *Transportation Research Part F: Traffic Psychology Behaviour*, vol. 65, pp. 620–629, 2018.
- [5] A. Williamson, D. A. Lombardi, S. Folkard, J. Stutts, T. K. Courtney, and J. L. Connor, "The link between fatigue and safety," *Accident Analysis & Prevention*, vol. 43, no. 2, pp. 498–515, 2011.
- [6] NHTSA, *A Sleep at the Wheel: A National Compendium of Efforts to Eliminate Drowsy Driving*, U.S. Department of Transportation, Washington, DC, USA, 2017.
- [7] Fatigue, 2017, <https://ec.europa.eu/transport/roadsafety/>.
- [8] M. L. Lee, M. E. Howard, W. J. Horrey et al., "High risk of near-crash driving events following night-shift work," *Proceedings of the National Academy of Sciences*, vol. 113, no. 1, pp. 176–181, 2015.
- [9] D. Dawson, A. C. Reynolds, H. P. A. Van Dongen, and M. J. W. Thomas, "Determining the likelihood that fatigue was present in a road accident: a theoretical review and suggested accident taxonomy," *Sleep Medicine Reviews*, vol. 42, pp. 202–210, 2018.
- [10] A. Anund, G. Kecklund, B. Peters, and T. Åkerstedt, "Driver sleepiness and individual differences in preferences for countermeasures," *Journal of Sleep Research*, vol. 17, no. 1, pp. 16–22, 2008.
- [11] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: recent insights and future challenges," *Journal of Safety Research*, vol. 40, no. 4, pp. 239–245, 2009.
- [12] G. Sikander and S. Anwar, "Driver fatigue detection systems: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2339–2352, 2018.
- [13] Q. He, W. Li, X. Fan, and Z. Fei, "Driver fatigue evaluation model with integration of multi-indicators based on dynamic Bayesian network," *IET Intelligent Transport Systems*, vol. 9, no. 5, pp. 547–554, 2014.
- [14] M. Doudou, A. Bouabdallah, and V. Berge-Cherfaoui, "Driver drowsiness measurement technologies: current research, market solutions, and challenges," *International Journal of Intelligent Transportation Systems Research*, vol. 18, no. 2, pp. 297–319, 2019.
- [15] J. Vicente, P. Laguna, A. Bartra, and R. Bailón, "Drowsiness detection using heart rate variability," *Medical, & Biological Engineering & Computing*, vol. 54, no. 6, pp. 927–937, 2016.
- [16] F. You, Y. Gong, H. Tu, J. Liang, and H. Wang, "A fatigue driving detection algorithm based on facial motion information entropy," *Journal of Advanced Transportation*, vol. 2020, Article ID 8851485, 17 pages, 2020.
- [17] Z. Li, S. 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, p. 495, 2017.

- [18] H. Zhang, C. Wu, Z. Huang, X. Yan, and T. Z. Qiu, "Sensitivity of lane position and steering angle measurements to driver fatigue," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2585, no. 1, pp. 67–76, 2016.
- [19] A. S. Zandi, A. Quddus, L. Prest, and F. J. E. Comeau, "Non-intrusive detection of drowsy driving based on eye tracking data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 6, pp. 247–257, 2019.
- [20] G. Sikander and S. Anwar, "Driver fatigue detection systems: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2339–2352, 2019.
- [21] A. Němcová, V. Svozilová, and K. Bucsuházy, "Multimodal features for detection of driver stress and fatigue: review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3214–3233, 2020.
- [22] S. Soares, S. Ferreira, and A. Couto, "Driving simulator experiments to study drowsiness: a systematic review," *Traffic Injury Prevention*, vol. 1, pp. 1–9, 2020.
- [23] Z. Li, S. Li, R. Li, B. Cheng, and J. Shi, "Driver fatigue detection using approximate entropic of steering wheel angle from real driving data," *International Journal of Robotics and Automation*, vol. 32, no. 3, pp. 291–298, 2017.
- [24] T. Wakita, "Driver identification using driving behavior signals," *IEICE-Transactions on Info and Systems*, vol. E89-D, no. 3, pp. 1188–1194, 2006.
- [25] R. Rossi, M. Gastaldi, and G. Gechele, "Analysis of driver task-related fatigue using driving simulator experiments," *Procedia-Social and Behavioral Sciences*, vol. 20, pp. 666–675, 2011.
- [26] K. Henni, N. Mezghani, C. Gouin-Vallerand, P. Ruer, Y. Ouakrim, and É. Vallières, "Feature selection for driving fatigue characterization and detection using visual- and signal-based sensors," *Applied Informatics*, vol. 5, no. 1, pp. 1–15, 2018.
- [27] R. Feng, G. Zhang, and B. Cheng, "An on-board system for detecting driver drowsiness based on multi-sensor data fusion using Dempster-Shafer theory," in *Proceedings of the International Conference on Networking*, Valencia, Spain, 2009.
- [28] A. D. McDonald, C. Schwarz, J. D. Lee, and T. L. Brown, "Real-time detection of drowsiness related lane departures using steering wheel angle," *Proceedings of the Human Factors and Ergonomics Society-Annual Meeting*, vol. 56, no. 1, pp. 2201–2205, 2012.
- [29] Q. Cheng, W. Wang, X. Jiang, S. Hou, and Y. Qin, "Assessment of driver mental fatigue using facial landmarks," *IEEE Access*, vol. 7, pp. 150423–150434, 2019.
- [30] X. Zhang, B. Cheng, and R. Feng, "Real-time detection of driver drowsiness based on steering performance," *Journal of Tsinghua University*, vol. 7, pp. 1072–1076, 2010.
- [31] E. C. Chua, S. C. Yeo, I. T. Lee et al., "Individual differences in physiologic measures are stable across repeated exposures to total sleep deprivation," *Physiological reports*, vol. 2, no. 9, pp. 1–16, 2014.
- [32] A. Bouhoute, R. Oucheikh, K. Boubouh, and I. Berrada, "Advanced driving behavior analytics for an improved safety assessment and driver fingerprinting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2171–2184, 2019.
- [33] S. Ezzini, I. Berrada, and M. Ghogho, "Who is behind the wheel? Driver identification and fingerprinting," *Journal of Big Data*, vol. 5, no. 1, 2018.
- [34] E. Miro, T. Alex, K. Karl, and T. Kohno, "Automobile driver fingerprinting," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 1, pp. 34–50, 2016.
- [35] I. d. Campo, R. Finker, M. V. Martínez, J. Echanobe, and F. Doctor, "A real-time driver identification system based on artificial neural networks and cepstral analysis," in *Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 1848–1855, 2014.
- [36] Y. Xun, J. Liu, N. Kato, Y. Fang, and Y. Zhang, "Automobile driver fingerprinting: a new machine learning based authentication scheme," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1417–1426, 2020.
- [37] M. Ingre, T. Akerstedt, B. Peters, A. Anund, and G. Kecklund, "Subjective sleepiness, simulated driving performance and blink duration: examining individual differences," *Journal of Sleep Research*, vol. 15, no. 1, pp. 47–53, 2006.
- [38] P. Thiffault and J. Bergeron, "Fatigue and individual differences in monotonous simulated driving," *Personality and Individual Differences*, vol. 34, no. 1, pp. 159–176, 2003.
- [39] X. Wang and C. Xu, "Driver drowsiness detection based on non-intrusive metrics considering individual specifics," *Accident Analysis & Prevention*, vol. 95, no. part B, pp. 350–357, 2016.
- [40] W. B. Verwey and D. M. Zaidel, "Predicting drowsiness accidents from personal attributes, eye blinks and ongoing driving behaviour," *Personality and Individual Differences*, vol. 28, no. 1, pp. 123–142, 2000.
- [41] R. Yan, C. Wu, and Y. Wang, "Exploration and evaluation of individual difference to driving fatigue for high-speed railway: a parametric SVM model based on multidimensional visual cue," *IET Intelligent Transport Systems*, vol. 12, no. 6, pp. 504–512, 2018.
- [42] M. Ingre, T. Akerstedt, B. Peters, A. Anund, G. Kecklund, and A. Pickles, "Subjective sleepiness and accident risk avoiding the ecological fallacy," *Journal of Sleep Research*, vol. 15, no. 2, pp. 142–148, 2010.
- [43] J. Jo, S. J. Lee, K. R. Park, I.-J. Kim, and J. Kim, "Detecting driver drowsiness using feature-level fusion and user-specific classification," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1139–1152, 2014.
- [44] C. S. Silveira, J. S. Cardoso, A. L. Lourenço, and C. Ahlström, "Importance of subject-dependent classification and imbalanced distributions in driver sleepiness detection in realistic conditions," *IET Intelligent Transport Systems*, vol. 13, no. 2, pp. 347–355, 2018.
- [45] A. Persson, H. Jonasson, I. Fredriksson, U. Wiklund, and C. Ahlstrom, "Heart rate variability for classification of alert versus sleep deprived drivers in real road driving conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–10, 2020.
- [46] F. You, X. Li, Y. Gong, H. Wang, and H. Li, "A real-time driving drowsiness detection algorithm with individual differences consideration," *IEEE Access*, vol. 7, pp. 179396–179408, 2019.
- [47] W. Chu, C. Wu, and H. Zhang, "Driver behavior model and its application in driver fatigue identification," *China Safety Science Journal*, vol. 28, no. 6, pp. 43–48, 2018.
- [48] C. J. d. Naurois, C. Bourdin, C. Bougard, and J.-L. Vercher, "Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness," *Accident Analysis Prevention*, vol. 121, pp. 118–128, 2018.
- [49] Y. Liang, W. J. Horrey, M. E. Howard et al., "Prediction of drowsiness events in night shift workers during morning



- driving,” *Accident Analysis & Prevention*, vol. 126, pp. 105–114, 2019.
- [50] A. Å. Miley, G. Kecklund, and T. Åkerstedt, “Comparing two versions of the Karolinska sleepiness scale (KSS),” *Sleep and Biological Rhythms*, vol. 14, no. 3, pp. 257–260, 2016.
- [51] H. Singh and A. Kathuria, “Analyzing driver behavior under naturalistic driving conditions: a review,” *Accident Analysis & Prevention*, vol. 150, Article ID 105908, 2021.
- [52] T. a. Åkerstedt, B. A. C. Anund, J. b. Axelsson, and G. Kecklund, “Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function,” *Journal of Sleep Research*, vol. 3, pp. 240–252, 2014.
- [53] T. Moriyama and Y. Maesono, “Smoothed alternatives of the two-sample median and Wilcoxon’s rank sum tests,” *Statistics*, vol. 52, no. 5, pp. 1096–1115, 2018.
- [54] L. Jin, K. Li, Q. Niu, and L. L. Gao, “A new method for detecting driver fatigue using steering performance,” *Journal of Transport Information and Safety*, vol. 32, no. 5, pp. 103–107, 2014.
- [55] K. Dejong, *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*, University of Michigan, Arbor, MI, USA, 1975.
- [56] B. Tavakkol, M. K. Jeong, and S. L. Albin, “Measures of scatter and fisher discriminant analysis for uncertain data,” *IEEE Transactions on Systems, Man, Cybernetics: Systems*, vol. 51, no. 3, pp. 1690–1703, 2019.