IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 29, 2020, accepted January 26, 2021, date of publication January 29, 2021, date of current version February 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3055551

XGBoost Algorithm-Based Monitoring Model for Urban Driving Stress: Combining Driving Behaviour, Driving Environment, and Route Familiarity

YUE LU, XINSHA FU, ENQIANG GUO, AND FENG TANG
School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China
Corresponding author: Yue Lu (201810101656@mail.scut.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 51778242 and Grant 51978283.

ABSTRACT Stress is considered by many studies to affect traffic safety, and many researchers have attempted to monitor the dynamics of driving stress. Previous research has relied excessively on the positive effects of psychological indicators to improve the accuracy of stress monitoring models. However, psychological data collection sensors have not been widely used in conventional vehicles, which makes it impossible to apply the results of that research to actual driving tasks on a daily basis, even if the accuracy is high. This study designs a real driving task to extract data and proposes a driver's driving stress monitoring model based on driving behaviour, driving environment, and route familiarity. The driving behaviour is described by the speed and acceleration of the vehicle, and the driving environment is quantified by a dilated residual networks (DRN) model thazt divides the video image from the full region into subregions according to the distribution of the driver's attention. Based on the psychological data and driver stress inventory (DSI) results, the study used a K-means 3D cluster analysis to obtain the evaluation method of driving stress and constructed an extreme gradient boosting (XGBoost) model to monitor driving stress. Comparisons of performance with other models show that the XGBoost model significantly outperforms the other three mainstream machine learning algorithms and exceeds most traditional models without the use of psychological data. The model's performance indicators, accuracy, sensitivity, and precision, reached 91.18%–93.25%, 84.13%–89.37%, and 90.25%–91.34%, respectively. The study also summarises the ranking of effects of different scene elements on driving stress for each visual field. The results could make it possible to apply stress monitoring on a large scale to real driving situations, providing urban designers with advice on how to reduce driver stress and directing their attention to those visual areas and visual scene elements that have a higher impact on driving stress and need improvement.

INDEX TERMS Driving stress, monitoring models, driving behaviour, driving environment, machine learning, route familiarity.

I. INTRODUCTION

As the scale of construction and the number of cars in cities increase, there is an urgent need to address the issue of urban traffic safety. Driver stress is increasingly recognised as a significant cause of accidents [1], with research from the Virginia Tech Transportation Institute showing that stress increases the risk of accident by a factor of 10 [2]. A growing number of researchers are attempting to dynamically monitor driver stress on the road, and psychological, driving behaviour, and environmental data are often used in their research [3]–[5]. Changes in psychological indicators are the most direct manifestation of changes in driving stress conditions. Previous research in stress detection has relied excessively on the positive effects of psychological indicators on model performance, and physiological data are obtained from sensors attached to the driver's body. These sensors are not widely used in conventional automotive assistance

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

systems owing to their uncomfortable and expensive nature [1]. This is why previous research on driving stress monitoring models, even with high accuracy, is still not applicable to everyday driving.

There are many sources of driver stress while driving, such as telephone calls and conversations in the car. The most important and controllable influences are driving behaviour and changes in the driving environment. Many researchers have identified driving behaviour as an important factor influencing driving stress and have quantified it using steering wheel rotation, car speed, and head movements [6]. Regarding the driving environment, owing to the lack of methods to quantify this factor that directly affects the driver's visual perception, it is often qualitatively analysed or even ignored [7]. Thanks to the application of deep learning in computer vision, its powerful feature extraction capabilities help us better understand the semantic information of each pixel in an image, providing a means to quantify the driving environment [8]. In addition, route familiarity has also been shown to affect driving stress, with unfamiliar routes increasing psychological pressure and tension for drivers, and familiar routes leading to more frequent risky driving behaviour and thus increased driving stress [9], [54].

Using experimental data from a real driving task, this study attempts to move away from the reliance of driving stress monitoring models on psychological data and to develop a real-time monitoring model of driving stress based only on the driving behaviour, driving environment, and route familiarity. The experimental data consisted of 1022 sets of observations, including physiological data, vehicle operation data, driving environment data, and Driver Stress Inventory (DSI) data, obtained by 23 experiment participants who drove seven consecutive trials on six widely varying and unfamiliar routes. The physiological and DSI data were used to establish the evaluation method of driving stress level using the K-means 3D cluster analysis, and the vehicle operating speed and acceleration were used to describe the driver's driving behaviour. Based on the dilated residual networks (DRN) model, the driving videos were used to describe the changes in the driving environment according to the distribution of the driver's attention between the full regional segmentation and the subregional segmentation. The study used the extreme gradient boosting (XGBoost) model to establish a driver's driving stress monitoring model and to explore the ranking of the effects of different scene elements on driving stress in each visual area. The contributions of this study are as follows.

1) We propose a new model for monitoring the driving stress based on the driving behaviour and driving environment, which achieves ideal performance without the use of psychological data. The results of this research can be used to monitor the driver's driving stress using conventional vehicle equipment, which makes it possible to apply stress monitoring to daily driving on a large scale.
2) The study not only ranked the different regions within the visual field in terms of their impact on driving stress, but also summarised the ranking of the impacts of different scene elements on driving stress for each region. We believe that the results of the study can provide urban designers with recommendations for reducing driver stress and direct their attention to visual areas and scene elements that have a higher impact on driving stress and need improvement.

The main structure of this paper is as follows. In Section 2, we review the research works related to this study. In Section 3, we describe the experimental design and feature extraction of this study. In Section 4, we describe the process of developing the XGBoost-based driving stress monitoring model. In Section 5, we include the results of this study, present the performance of the model and compare it with other related studies, and finally explore the importance of different variables on the model. In Sections 6 and 7, we discuss the conclusions and implications of this study and identify the shortcomings.

II. RELATED WORK

A. METHODS FOR MEASURING DRIVING STRESS

Stress is an important factor in driving because it affects a driver's performance, and stress is considered as one of the most important causes of car accidents [10].Stress in psychology represents a state of feeling in which humans are unable to respond normally to external stimuli, and researchers distinguish between eustress and distress. Eustress refers to a stress that allows a creature to produce a more beneficial state, such as vigilance [11]. However, more researchers consider stress as distress and explore the negative effects it brings to people. In terms of traffic, driving requires constant dynamic attention to the driving task and the ability to identify potential hazards and monitor changes in a complex environment, and driving stress can affect traffic safety by influencing the driving task as described above [12]. In fact, many studies have confirmed that driving stress affects traffic safety. Excessive stress can disrupt driver performance and reduce driving safety [13], and accident reports have shown that stress significantly increases the probability of accidents.As researchers have realised the impact of driving stress on traffic safety, they have begun to explore how to measure driving stress comprehensively.Stress is often caused by unpleasant external interventions and is influenced by the participant themselves and their environment, so researchs generally measure driving stress in four ways [14].

1) PSYCHOLOGICAL EVALUATION

Psychological evaluation is measured mostly in the form of questionnaires [15], [16], the Driving Behaviour Questionnaire(DBQ) and the Driver Stress Inventory(DSI) are two of the most commonly used methods to measure psychological states [17].The DBQ evaluates the psychological state of the driver by measuring three types of abnormal

driving behaviours: violations, errors, and mistakes [12], but the DBQ does not measure the emotional response to the combined measure of driving.The revised DSI, based on the DBQ, assesses the driver's stress responses while driving on five different dimensions and obtains good results, including aggressiveness, dislike of driving, hazard monitoring, fatigue tendency, and thrill seeking. As researchers have found similar influences on driver stress in different cultures [18], [19], the DSI has been translated into multiple languages based on language, traffic rules, and driving habits in different countries [20].

2) INTERNAL PHYSIOLOGICAL RESPONSES

Stress is essentially regulated by the autonomic nervous system through the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS) in response to physiological stress, thus adrenaline and cortisol levels can be used as two main physiological indicators to detect stress [23], but epinephrine is not commonly used because of its invasive method of acquisition. Other psychological data are also commonly used for stress measurement, such as Electrocardiogram(ECG), Electrodermal Activity(EDA), Respiration Activity(RSP), Electromyography(EMG), Skin Temperature, and Pupillary Dilation. Salivary cortisol levels have been shown to measure driver stress-induced sympathetic nerve activity as well as the stress produced [39].Two parameters in the ECG, heart rate(HR) and heart rate variability(HRV), can indicate fluctuations in the ANS and have been widely used for the measurement of driver stress levels [21], [22]. At the same time, EDA [23], [43], RSP [24], EMG [25], Skin Temperature [26], and Pupillary Dilationhave [27] also all been proven in research to be effective in measuring driver stress levels.

3) EXTERNAL PHYSICAL RESPONSE

In addition to involuntary physiological responses to stress, external and voluntary behavioural responses such as facial, verbal, and driving behaviours can also measure a driver's stress status. Some studies have aimed to identify drivers' stressful situations by identifying their facial reactions under negative emotions [28] and analyzing the sound waveforms of their verbal communication [29]. More researchers are opting to use vehicle dynamics data to represent drivers' driving behaviour and determine their psychological state, but the sources of data are diverse. Vehicle speed, acceleration, braking frequency and steering wheel angle are common vehicle dynamics data [23], [30], and even Global Positioning System(GPS) signals have been used by researchers to extract vehicle trajectories and describe their changes in dynamics [31]. Regardless of the data source, vehicle dynamics data has been shown to better predict driver stress conditions when added to research [32].

4) EXTERNAL ENVIRONMENT

Driver stress can also be affected by the external environment, such as traffic jams, heavy rain, night driving, etc [6], [33]. The parameters of the external environment can be divided into three categories: (1) Weather conditions [34]; (2) Driver visibility [35]; (3) Roadscape; (4) Driving routes [9]. However, due to the lack of accurate and efficient methods to quantify the driving environment, the external environment is an important factor that has not been extensively studied due to the complexity of statistical and analytical methods, for example, geographic databases have been used to provide information on road conditions, road types, intersection types, number of lanes, traffic signs and road curvature to describe the driving environment, the complexity of which can be imagined [49].

B. RESEARCH ON THE DRIVING STRESS MONITORING MODEL

With the refinement of driving stress measurement methods, researchers have attempted to monitor drivers' driving stress in real time to improve road safety. The research on driver pressure monitoring models falls into two main categories. The first approach is the construction of traditional statistical models to classify driver stress levels [36]–[38], which requires prior exploration of the association between numerous characteristics extracted from physiological signals and stress status, and is more difficult but less dependent on the amount of data. The second approach is the use of advanced techniques such as machine learning to construct models, which has become common with the development of data extraction and analysis techniques [6].The implementation of this method relies on a large amount of data, but the accuracy of the model improves significantly over the former.

The source of data for the study of stress monitoring models limits the accuracy and feasibility of the study to some extent, with psychological evaluations generally using cluster analysis to obtain taxonomies of stress levels [17], while internal physiological responses, external physical responses and the external environment are used to construct monitoring models. Since physiological responses are involuntary responses of the driver to stress, more studies have used physiological indicators as the main source of data due to their high accuracy and easy data extraction [40], and monitoring models rely on easily extractable and continuous physiological data such as skin conductance response(SCR), HR, blood pressure and EMG. Urbano *et al.* used a linear discriminant classifier to detect two levels of stress in drivers using EDA and SCR, and the model achieved accuracy of 81%–97% [41]. Another study proposed a three-level driver stress (low, medium, and high) detection method based on a cyclic neural network using photoplethysmogram (PPG), EDA. and RSP signals as eigenvalues and achieved an accuracy of 89.23% [40]. Rigas *et al.* used ECG, EDA, RSP signals and vehicle operation data to detect the two stress levels of the driver by using a plain Bayesian classifier, and the model achieved an accuracy of 96% [32]. Another study also used electrocardiograph (ECG), EDA, and RSP signals and vehicle operation data (steering wheel,

car speed, etc.), from which 42 features were selected to create a model with an accuracy of over 90% [23]. Rigas *et al.* also developed a driver stress monitoring model with an accuracy of 86% through a support vector machine (SVM) classifier using data from three sources: physiological signals (ECG, EDA and RSP), head movements, and the environment [6].

Although research on driver stress monitoring models have achieved ideal performance, the over-reliance on the positive effects that psychological indicators can have on improving accuracy and the fact that physiological sensors are not standard in automotive assistance systems has led to stress monitoring not being used on a large scale in the automotive industry.Automakers are currently trying to add sensors that capture physiological data into car assistance systems. Ford is working on a technology that could restrict driver cell phone use when the system detects congestion and stressful driving conditions [42]. Automakers can add heart rate and body temperature sensors to steering wheels, and respiratory rate sensors to seatbelts. In addition, the BMW Group, in cooperation with other research institutes, has developed a sensor system that is integrated in the steering wheel to monitor physiological data such as the driver's heart rate, skin conductivity, and blood oxygen saturation. The aim is to initiate appropriate measures when the system detects a stress condition or uncomfortable situation for the driver, such as blocking telephone access or turning off radio sounds [41]. This study attempts to rely on existing data acquisition devices in automobiles to monitor the driver's driving stress in terms of both driving behaviour and driving environment, which could make it possible to apply stress monitoring to daily driving on a large scale.

III. DATA AND FEATURE EXTRACTION

A. REAL DRIVING TASKS

1) PARTICIPANTS

In total, 9 female and 14 male university students (mean age = 22.6 years, SD = 1.7) from South China University of Technology (SCUT) were recruited through course internships and campus recruitment (email, social media) and paid to participate in the experiment. To ensure the scientific validity of the experimental data, all participants were required to have two or more years of driving experience, have recently maintained continuous driving status, and complete a questionnaire and sign an informed consent form prior to the experiment. The questionnaire included several questions affecting the experimental data such as history of heart disease, recent drug use, and recent intoxication, and the informed consent form was completed and reported to the local ethics committee, which included the purpose, methods, and risks of the experiment. The participants must also ensure that they hold a Chinese C1 driver's license, have normal or corrected visual acuity of 0.8 and above, and have no experience driving on the experimental roads.

2) EQUIPMENT

The data obtained during the real driving task consisted of the driver's physiological signals, the vehicle's operating status, and the data of the driving environment. Three instruments were used to ensure the comprehensiveness and accuracy of the data.

1) **BIOPACMP160**, which includes five data acquisition channels (ECG, EDA, and tri-axial acceleration (X/Y/Z)) to record the ECG, EDA, and vehicle acceleration data of the driver during driving. Among them, the ECG sensor is fixed on the front chest and right and left waist to collect the ECG signal, the EDA sensor is fixed on the forefinger and middle finger of the non-useful hand to collect the electrodermal signal, and the three-axis acceleration sensor is fixed on the horizontal position of the vehicle roof and calibrated in advance to collect the acceleration data when the vehicle is running. The sampling frequency of the BIOPACMP160 was set to 2000 samples/second, the signal acquisition interval of the ECG was adjusted to $+/-$ 10 mv, and the maximum response of the EDA was set to 50 microsiemens.
2) **YOUJIA i-box**, which can be connected to the test vehicle's on-board diagnostics (OBD) to acquire more than 200 real-time operating parameters, such as vehicle speed, fuel consumption, and water temperature, while driving (`http://www.gooddriver.cn/#/hobd`). In this study, only the operating speed is extracted to quantify the driver's driving behaviour, and the data sampling frequency is set to 2 Hz.
3) **Driving recorder**, which can record in real time the changing situation of the outside driving environment when a vehicle is running, and thus this study uses the driving recorder during operation of the test vehicle. The image acquisition resolution of the vehicle recorder is set to 2560 $\times$ 1600 to ensure the accuracy of image feature extraction.

3) PROCEDURE

This pilot of this study was conducted in Shenzhen, China, from May to December 2019. To eliminate the specificity of the experimental data, the experimental sites were selected in six regions with different community landscapes and road network conditions in Luohu District, Futian District, and Nanshan District, Shenzhen. Regarding the urban development features in Shenzhen, Luohu District, Futian District, and Nanshan District correspond to old town, CBD, and suburban areas, respectively. A closed experimental route is defined in each region and the road sections and intersections are numbered in advance to facilitate later data processing, and it is forbidden to inform the participant of the defined route before the experiment. The six selected experimental routes contained a total of 458 sections and intersections, with the intersections including 52 left turns, 59 right turns,

and 121 straight segments, and each route involved a travel time of approximately 1 h.

The right-hand driving rule was adopted in the experiment. To ensure the successful extraction of image features later, the experiment was carried out under good weather conditions, and the participants were required to wear the apparatus for test driving on a non-experimental route to reduce the discomfort of wearing the apparatus before the formal experiment. The experimental procedure consisted of deploying the experimental equipment, pre-trip data calibration, driving, removing the experimental equipment, and a final questionnaire survey. To ensure unfamiliarity with the experimental route, the participants were informed of their route while wearing the device. The participants were also asked to sit still and relax in their seats for more than ten minutes to measure the initial values of each physiological indicator in the driver's state of calm. During the experiment, participants were required to obey traffic rules, to eliminate the occurrence of non-study variables influencing driver perception of driving, and to ensure that they had a good field of vision while driving. In addition, the experiment time was set to 8:30 a.m.–5:00 p.m. on weekdays to ensure a similar type of traffic during the experiment. Participants were asked to drive more than 7 trials on each experimental route, resulting in a total of 45 trials each, for a total of 1035 sets of data. ECG data were missing or missing in 13 of the total of trials, and thus only 1022 sets of data were used in the final analysis.

It is important to note that as this study involves the collection of physiological data, the local ethics committee was informed of the study beforehand. However, considering that no invasive equipment was used, and the safety of the participants was not compromised, the ethics committee only asked for advance notice of the experiment details and informed consent.

4) DRIVER STRESS INVENTORY

To investigate the driver's driving stress, it is first necessary to find a way to quantify it. Researchers have been exploring ways to measure driving stress. Initially, a table called the DBQ was developed to evaluate the dimensions of driver stress [12], but it did not provide a comprehensive measure of driving emotions. The DSI, which is based on the DBQ, has been revised to evaluate driver stress vulnerability in five different dimensions, and has been very effective. As researchers have found similar influences of driver stress in different cultures [18], [19], the DSI has been translated into multiple languages based on language, traffic rules, and driving habits in different countries [20]. The Chinese researchers not only revised the form's provisions to fit China's situation, but also simplified the original 48 items to 27. Cronbach's alpha coefficient was used to perform reliability tests on different options, and the Chinese version of the DSI also included five factors: aggression (6 items;$\alpha = 0.7$); dislike of driving (6 items; $\alpha = 0.63$); hazard monitoring (5 items; $\alpha = 0.67$); proneness to fatigue (5 items; $\alpha = 0.75$); and thrill seeking (5 items; $\alpha = 0.73$) The stress inventory used an 11-point Likert scale to reflect participants' agreement with each item, with scores ranging from 0 (strongly disagree) to 10 (strongly agree), and the Chinese version of the driver stress inventory was found to be reliable and valid [44].

In this study, a Chinese version of the DSI was used to quantify the participants' stress status, and the participants were asked to review their stress status on each trial after the experiment to fill out the DSI for stress self-assessment. Through post-test communication, the participants revealed that they encountered almost all the emotional changes involved in daily driving during the experiment, proving the completeness and generalisability of the experimental data.

B. FEATURE EXTRACTION

1) PHYSIOLOGICAL DATA

a: ELECTROCARDIOGRAPH

ECG is the electrical signal generated by recording each cardiac cycle of the heart, which has been shown to be the most effective way to evaluate a driver's stress level, and an increasing number of studies of ECG signals in drivers have shown that selecting an appropriate method for quantifying the eigenvalue is critical. This study was conducted by extracting eigenvalues of HR from ECG signals, one of the most commonly used psychological indicators, which is often used to quantify a person's level of arousal and mental effort [45]. As the range of HRs varies between individuals, the study uses the relative HR, $HR_{relative}$, as an eigenvalue, which is quantified by the following formula:

$$HR_{relative} = \frac{HR}{HR_{base}} \tag{1}$$

where HR_{base} is the basal heart rate obtained when the participant was asked to sit still for ten minutes at the beginning of each experiment.

b: ELECTRODERMAL ACTIVITY

EDA is a characteristic of the human body used to describe the changes in skin resistance or electrical conductivity due to changes in the function of the skin sweat glands. Emotional changes will lead to changes in sweat gland secretion, which in turn will lead to changes in skin conductivity resulting in changes in the skin electrical index. Electrodermal activity is strongly correlated with mood, attention, and arousal, and researchers have widely used the electrodermal index to describe driver stress levels [46], [47]. Studies have shown that the skin conductance response should be chosen as a parameter to study when the level of psychological change is low, and the present study is a low-strain research on the change of driver stress level under normal driving conditions [48].

However, owing to the existence of a signal baseline, the skin conductance response does not fully represent the driver's psychological condition, and the threshold of skin level is not exactly the same for different drivers. Therefore, the SCR was chosen as the eigenvalue of EDA. The SCR removes the signal baseline on the basis of the original signal

of EDA and only considers the relative amplitude of the signal, which can better represent the change of stress status caused by external stimuli.

2) DATA ON OPERATIONAL STATUS

Driving behaviour has been shown to affect the driver's driving stress, including operating speed and acceleration [31]. The study attempts to characterise the driver's driving behaviour through the vehicle's operating status data, which are mainly running speed and three-axis acceleration, where the three-axis acceleration is obtained from the TSD109C2 sensor of BIOPACMP160 and the running speed is obtained from the vehicle's OBD system through the YOUJIA i-box. The TSD109C2 sensor needs to be calibrated in the direction (X/Y/Z) and fixed in the horizontal position inside the vehicle before the experiment, and uncontrollable factors such as engine vibration and uneven road surface may lead to interference signals in the original signal. Due to the small amplitude and low frequency of the interfering signals, a finite impulse response (FIR) low-pass filter based on a Blackman −61 dB window was used to filter out interfering signals with frequencies greater than 5 Hz to obtain the actual acceleration signal.

Owing to the small slope of urban roads, for the Z-axis (vertical component) of the three-axis acceleration, the interference signal from vehicle bumps caused by an uneven road surface is more obvious and difficult to filter compared to the original signal, and thus only the X- and Y-axis accelerations are considered in the study. In addition, the initial value of acceleration after calibration of the TSD109C2 sensor is not zero and the sensor needs to be recalibrated before each experiment, so the study uses $a_{x/y}$ as the eigenvalue of the operating status, which is quantified by the following formula:

$$Acceleration_{x/y)} = \frac{Acceleration_{original(x/y)}}{Acceleration_{base(x/y)}} \tag{2}$$

where $Acceleration_{base(x/y)}$ is the acceleration of the vehicle at standstill after the sensor has been recalibrated.

3) DATA ON THE DRIVING ENVIRONMENT

The driving environment is the driver's most intuitive visual perception of the external environment, which can directly affect the driver's psychological burden, and there have been research attempts to find a deep learning-based driving environment quantification method. With the widespread application of deep learning in computer vision, its powerful image extraction capability has been used by a growing number of researchers for image processing. Image semantic segmentation is an important research direction in image processing, the task of which is to make a dense prediction of the pixel information in an image, and the process is also called 'pixel prediction' [50]. With the improvements in deep learning theory and numerical computing devices, convolutional neural networks have been rapidly developed, and they also provide a quantitative means for image semantic segmentation. However, traditional convolutional neural networks add downsampling and pooling layers to obtain more abstract image features, and the spatial resolution of the corresponding feature map is reduced. However, the accuracy of the image segmentation task for pixel prediction purposes depends on the size of the resolution of the output features. The process of downsampling and pooling in traditional convolutional neural networks results in the loss of information about the spatial structure of scene elements, which in turn reduces the accuracy of image semantic segmentation and is detrimental to the model's ability to identify small targets and relationships between targets.

TABLE 1. Network configuration diagram for the dilated residual networks model.

Name	Filters	Dilation	Output Size
Con1_x	$7\times7, 64, stride2$	1	112×112
Con2_x	$3\times3\ maxpool, stride2$	1	56×56
	$\begin{bmatrix}3\times3,\ 64\\3\times3,\ 64\end{bmatrix}\times2$		
Con3_x	$\begin{bmatrix}3\times3,\ 128\\3\times3,\ 128\end{bmatrix}\times2$	1	28×28
Con4_x	$\begin{bmatrix}3\times3,\ 256\\3\times3,\ 256\end{bmatrix}\times2$	2	28×28
Con5_x	$\begin{bmatrix}3\times3,\ 512\\3\times3,\ 512\end{bmatrix}\times2$	4	28×28
Con6_x	$\begin{bmatrix}3\times3,\ 512\\3\times3,\ 512\end{bmatrix}\times2$	2	14×14
Con7_x	$\begin{bmatrix}3\times3,\ 512\\3\times3,\ 512\end{bmatrix}\times2$	1	7×7
	$Averagepool, softmax$		1 ×1

In this study, the DRN model is used for semantic segmentation, and the detailed model network configuration is presented in Table 1. The model is based on Resnet and removes the downsampling top layers to maintain the spatial resolution of the feature map. In addition, the model replaces the downsampling layer with an expansion convolution and applies it rationally to the subsequent layers to ensure the resolution of the receiver field of the subsequent layers and reduce significantly the loss of image pixel details [51]. To ensure that image information is not lost in transmission and to prevent the memory consumption in the image segmentation task from exceeding hardware capacity, the DRN model sets up expansion convolution only at the fourth (Group4) and fifth (Group5) of the model's seven groups of convolution. Owing to the increase in model expansion rate, the larger spacing of pixel samples causes the model to lose local information and produce grid effects. The model adds Group 6 and Group 7 convolution at the end of the network to avoid over-expansion of the mesh. We trained the model using the Cityscapes dataset (`https://www.cityscapes-dataset.com/`), and the model was shown to have higher accuracy on the Cityscapes dataset than the ResNet-101 model. The image semantic segmentation flow is shown in Figure 1.

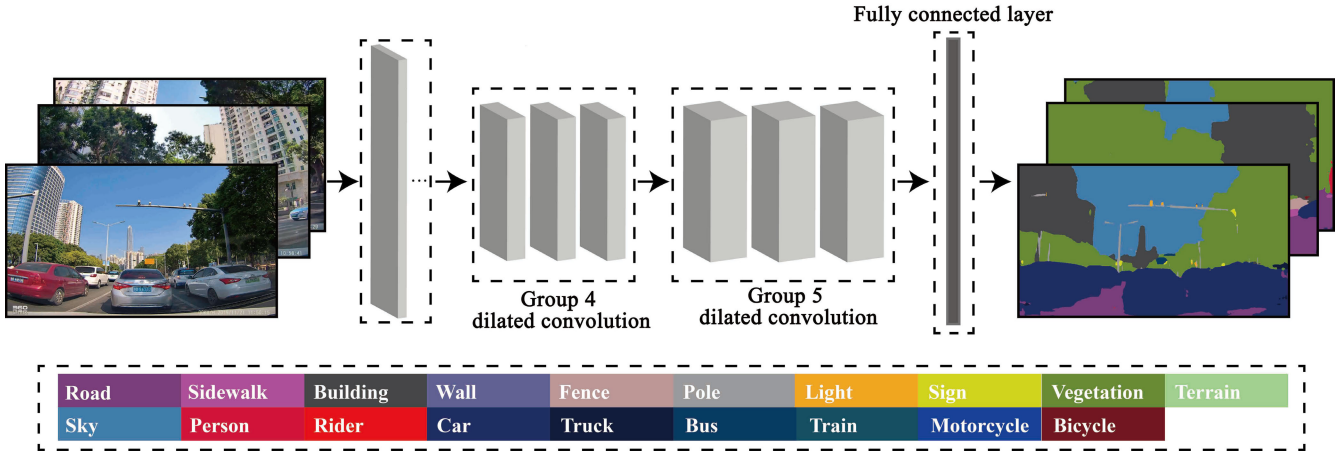


FIGURE 1. Image segmentation process based on dilated residual networks.

Two methods for extracting image information, full regional segmentation and subregional segmentation, are provided with the aim of finding the most appropriate representation of the driver's visual perception. Full regional segmentation analyses the influence of the information of different scene elements in the field of view on the driver's perception of the driving environment from the global perspective, which simplifies the computation but loses some feature information of the scene elements. It is well known that different categories of scene elements within different fields of view affect driver physiological indicators to different degrees. For example, the presence of pedestrian elements on the road and on the pavement causes driver stress to a completely different degree. The study proposes an information extraction method based on full regional segmentation for subregional segmentation. Thanks to the researcher's study of the distribution of attentional viewpoints in different regions of the visual range [35], and taking into account the size conversion between the viewpoint of the driving video and the driver's viewpoint, the study divides the picture into five regions, Up, Bottom, Left, Right, and Centre, according to the distribution of attention, as shown in Figure 2. Although this method can capture more information about the hidden details in the images, excessive segmentation can lead to more zero values for some of the eigenvalues, which is detrimental to the training and optimisation of the model. In either method, the study acquires eigenvalues of the driving environment from three sources.

a: PROPORTION OF VISUAL SCENE ELEMENTS

The whole driving environment is a combination of different scene elements, and changes in the proportion of scene elements are the most visible manifestation of changes in the driving environment. The proportion of different scene elements can intuitively affect the driving environment, and thus the driver's perception, while the position of different elements in the scene can also greatly influence the psychological changes of the driver. Therefore, the study takes the proportion of visual scene elements as a quantitative indicator

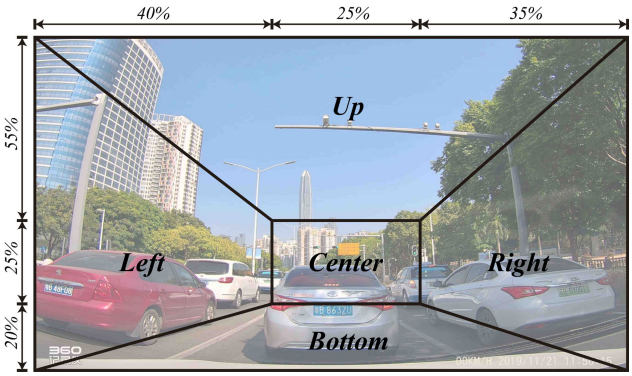


FIGURE 2. Segmentation of driving video images.

of the driving environment, with the following formula:

$$Proportion_{ij} = \frac{\sum_{s=1}^{n} pixel_{ijs}}{\sum_{j=1}^{19} \sum_{s=1}^{n} pixel_{ijs}} \tag{3}$$

where $pixel_{ij}(i = All, Up, Bottom, Left, Right, Centre; j = 1, 2, \ldots, 19)$ denotes the proportion of the j-th scene element in the i-th region and $pixel_{ijs}$ denotes the s-th pixel of the j-th scene element in the i-th region.

b: TIME SERIES CHANGE OF INFORMATION ABOUT THE PROPORTION OF VISUAL SCENE ELEMENTS

The driving environment changes continuously when the vehicle is moving, and the static data of visual scene elements only shows the distribution of visual scene elements in the driver's field of view at the current moment, but does not describe the temporal continuity well. As this study extracts pictures from the driving video at a frequency of 6 frames/s, we use the difference between the proportional information of the current frame and the average proportional information of the previous 6 frames to obtain the time series change of information of the proportion of visual scene elements, $DiffProportion_{ij}$, as a quantitative indicator to describe the continuity of the driving environment. The quantitative for-

mula is as follows:

$$DiffProportion_{ijt} = Proportion_{ijt} - \frac{\sum_{n-1}^{6} Proportion_{ij(t-n)}}{6} \tag{4}$$

where $DiffProportion_{ijt}(i = All, Up, Bottom, Left, Right, Centre; j = 1, 2, \ldots, 19)$ denotes the time series change in the proportion of the j-th scene element in the i-th region, and $Proportion_{ijt}$ denotes the proportion of the j-th scene element in the i-th region of the t-th frame picture.

c: NUMBER OF TYPES OF MOVEABLE SCENE ELEMENTS

Thanks to previous research on interference factors in driving, we understand that the level of information on different scene elements is different for drivers, and that moveable scene elements are more interfering compared to other elements [52], [53]. Five common movable scene elements exist in the 19 scene elements obtained by the image segmentation method based on the DRN model, including Person, Car, Bus, Motorcycle, and Bicycle, and the study adopts the number of types of movable scene elements A as the quantifiable indicator. To avoid a segmentation error rate, we quantified the indicator by the following formula:

$$Movetype_i = \sum_{j=r}^{7} movetype_{ir} \tag{5}$$

$$movetype_{ir} = \begin{cases} 0, & if\ Proportion_{ir} \leq 0.05 \\ 1, & if\ Proportion_{ir} > 0.05 \end{cases} \tag{6}$$

where $Movetype_i(i = All, Up, Bottom, Left, Right, Centre)$ is the number of types of moveable scene elements in the i-th region and $movetype_{ir}(i = 1, 2 \ldots, 5)$ is the judgment value of the r-th moveable scene element.

4) ROUTE FAMILIARITY

It has been shown that a driver's familiarity with the route affects his or her psychological state when driving [9], [54], Since participants had no driving experience on the test road prior to the experiment, their familiarity with the test route increased with the duration of the trial. Therefore, the study used the number of trial laps as the route familiarity and quantified the route familiarity using the following formula.

$$Routh\_Fam_i = Test\_Laps_i \tag{7}$$

where $Routh\_Fam_i(i = 1, 2, \ldots, 6)$ denotes is the route familiarity indicator for the i-th test route, and $Test\_Laps_i$ denotes the number of test laps of the i-th test route. The experiment required participants to drive seven consecutive laps to familiarise themselves with the test route, but in reality drivers do not drive continuously in order to familiarise themselves with the route, and discontinuous driving can lead to route forgetting, so in reality drivers will be slower to familiarise themselves with the route than in this test. This study assumes that the number of trials also reflects, to some extent, the degree of route familiarity, and therefore ignores the subtle differences between the two.

5) EVALUATION OF DRIVING STRESS LEVELS BASED ON K-MEANS 3D CLUSTERING ANALYSIS

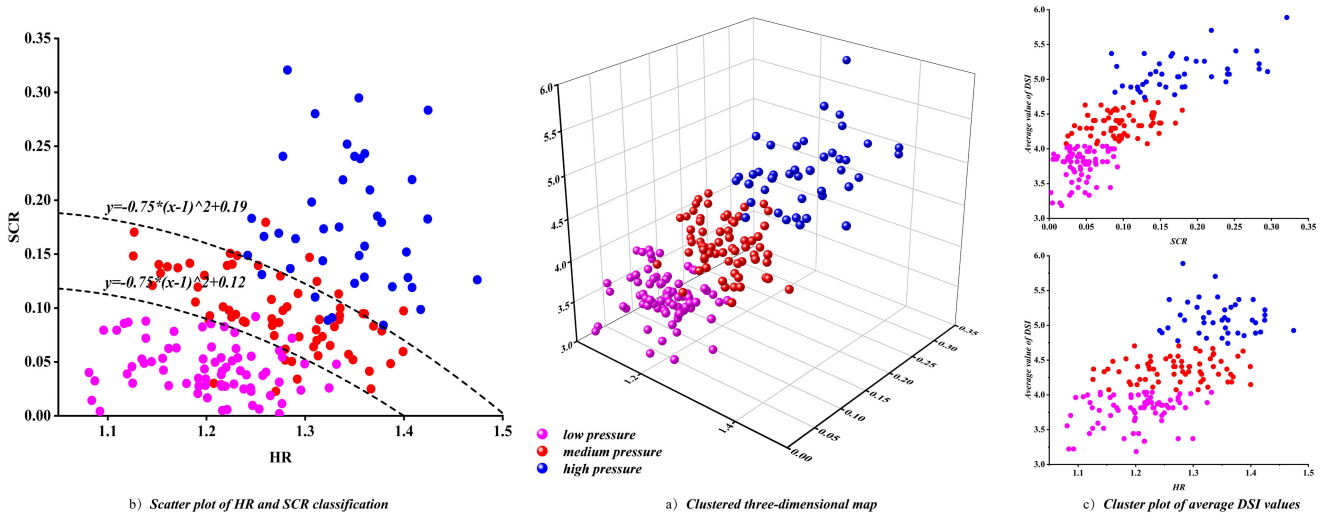
How to measure driver stress levels has long been a concern for researchers, and as previous studies have demonstrated a strong relationship between ECG, EDA, and driving stress [5], they have attempted to find ways to quantify the driver psychological status and stress levels, and hopefully evaluate the driver stress status based on physiological indicators. To determine the driver's stress status during the experiment, the driver was asked to complete the DSI after finishing the test by reviewing the driving video. The average value of the DSI, $DSI_{average}$, was used to indicate the driver's stress level on each trial. The study used the data from each trial as a sample dataset to explore how the stress status could be evaluated through physiological data. First, we collected ECG, EDA, and stress data in one trial and matched them to each other, with ECG data as the mean relative HR per trial, $HR_{relative}$, electrodermal data as the 85th percentile of the SCR value per trial, $SCR_{85\%}$(similar to $V_{85\%}$ for running speed), and stress data as the mean of all items in the DSI, $DSI_{average}$. In total, 200 sets of sample data were used to evaluate the method in the analysis, and the results of the statistical analysis of the three specific descriptions are listed in Table 2.

Second, considering that the sample size and data span for the psychological indicators were not large, and drawing on relevant previous studies, we attempted to classify the driving stress scale into high, medium and low by means of HR and SCR, and unsupervised learning cluster analysis was used for this study. However, the clustering analysis of $DSI_{average}$ only classifies the stress level from the perspective of the numerical value, but does not take into account the descriptive effect of HR and SCR on the stress. Thus, the study attempts to evaluate the driving stress and its level using the K-means 3D cluster analysis method. Figure 3a) shows a three-dimensional diagram of the cluster delineation results, with different colours representing clusters of three driving stress levels: high, medium, and low, respectively. Figures 3b) and 3c) present the three views of the three-dimensional diagram, respectively. The classification function for different stress levels on the HR–SCR viewpoint diagram can be obtained by regressing the boundary data, and the driving stress level of each experimental sample, $ClassifiLevel$, can be quantified by the equation 8, as shown at the bottom of the next page.

The statistical ranges for the description of different stress levels $HR_{relative}$, $SCR_{85\%}$, and $DSI_{average}$ after 3D cluster-ing are listed in the table.

IV. METHOD

This study proposes an XGBoost-based model of the driver's driving stress monitoring for classifying driving stress levels based on driving behaviour, driving environment, and route familiarity, and the model construction and training methods are shown in Figure 4.



b) *Scatter plot of HR and SCR classification* a) *Clustered three-dimensional map* c) *Cluster plot of average DSI values*

FIGURE 3. Graph of driving stress assessment results based on K-Means 3D clustering analysis.

TABLE 2. Statistical analysis table of descriptions of $HR_{relative}$, $SCR_{85\%}$, $DSI_{average}$, and their composition.

	$DSI_{average}$						$HR_{relative}$	$SCR_{85\%}$
	Aggression	**Dislike**	**Hazard**	**Fatigue**	**Thrill**	**Total**		
Items	6	5	5	5	6	27	–	–
Mean	4.52	5.35	5.77	3.33	2.81	4.3	1.260	0.094
Standard Deviation(SD)	0.67	0.74	0.85	0.87	0.56	0.55	0.083	0.065
Classification low	3.2–5.3	3.8–6.0	4–7.2	2–4.8	1.7–3.3	3.18–4.04	1.08–1.33	0.002–0.092
medium	4.0–5.7	4.6–6.8	4.8–7.6	2.0–5.0	2.0–4.0	4.07–4.70	1.125–1.399	0.093–0.179
high	4.3–6.5	5.0–7.0	6–7.6	2.0–7.0	2.7–4.3	4.74–5.89	1.241–1.474	0.169–0.320

A. EXTREME GRADIENT BOOSTING–XGBoost

In this study, we selected the XGBoost algorithm to monitor the driver's driving stress. XGBoost is an integrated algorithm based on a linear classifier or tree, which produces a strong classifier with better classification or regression results by integrating several weak classifiers. The traditional and most commonly used single classifiers, such as Decision Tree and Support Vector Machines (SVM), tend to improve classification performance by increasing the amount of training data. After a certain amount of training data, the increase in the amount of data is no longer significant to improve the accuracy of the model, and researchers have considered fusing multiple models from both Bagging and Boosting to improve the accuracy of the overall classifier. Among them, Gradient Boosting Decision Tree (GBDT) is a recent popular strong learner based on multiple decision trees and gradient boosting iterations, where the model achieves satisfactory performance by running a certain number of trees. However, when the dataset is large and the model is complex, the algorithm may require thousands of iterations to complete the training of the model and there is a large computational complexity.

XGBoost has become the most popular enhancement tree algorithm for the gradient boosting machine (GBM), which optimizes the objective function based on GBDT to ensure the accuracy of the model while controlling the complexity of the model, and it is widely used in industry owing to its high performance in classification and regression and its low requirements for feature engineering. Similar to the gradient boosting algorithm, XGBoost is based on an integrated learning approach consisting of classification and regression trees (CART) [55]. The model continuously adds and trains new trees in each iteration to fit the residuals of the predicted values of the previous decision tree and the sum of the predicted values of all previous decision trees, and finally sums the predicted values of all the decision trees together as the final result. It is a hierarchical method based on weighted extreme gradient boosting in the ECG heartbeat classification [56], [57].

$$ClassifiLevel = \begin{cases} High, & if \quad SCR_{85\%} + 0.75 \times (HR_{relative} - 1)^2 > 0.19 \\ Medium, & if \quad 0.19 \geq SCR_{85\%} + 0.75 \times (HR_{relative} - 1)^2 > 0.12 \\ Low, & if \quad SCR_{85\%} + 0.75 \times (HR_{relative} - 1)^2 \leq 0.12 \end{cases} \tag{8}$$

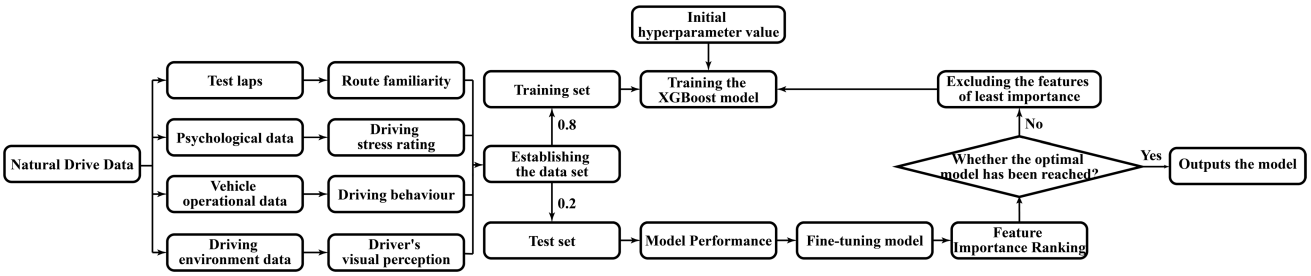


FIGURE 4. Method flow chart.

For an integrated learning model with K trees, if the input sample is x_i, the model is defined as follows:

$$\begin{aligned}\widehat{y_i} &= F_K(x_i) = F_{K-1}(x_i) + f_K(x_i) \\ &= \sum_{k=1}^{K} f_k(x_i), f_k \in F \text{ and } f_k(x_i) = \omega_{q(x)}\end{aligned} \tag{9}$$

where $f_x(x_i)$ denotes the mapping of the k-th tree, F_K is the set of K response functions f_x, $\omega_{q(x)}$ denotes the predicted value of the sample falling in this leaf node in this tree, and $q(x)$ denotes the number of leaf nodes in this tree. The objective function of XGBoost consists of two parts: the first is the loss function, which is used to evaluate the error between the predicted and true values of the model; the second is the regularisation function, which is used to control the complexity of the model, and the objective function of XGBoost is defined as follows:

$$Obj = \sum_i l(\widehat{y_i}, y_i) + \sum_k \Omega(f_k) \tag{10}$$

where l is a loss function that quantifies the difference between predicted values $\widehat{y_i}$ and label y_i. Ω is a regularisation function that measures complexity and tends to choose simple models to prevent overfitting of the model. The regularisation function is defined as follows:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|\omega_i\|^2 \tag{11}$$

where T denotes the number of leaf nodes and γT controls the complexity of the model by the number of leaf nodes and their coefficients. Here, the larger the value is, the larger the objective function. The second term is the regular term of $L2$, which controls the weight scores of the leaf nodes, and ω_i denotes the score of the i-th leaf. The objective function cannot be optimised in European space, and considered from an approximate perspective, the model prediction score for the t-th round sample x_i can be expressed as

$$\widehat{y_i}^{(t)} = \widehat{y_i}^{(t-1)} + f_t(x_i) \tag{12}$$

Thus, the objective function can be expressed as

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \widehat{y_i}^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{13}$$

where g_i is the first-order gradient statistic of the loss function and h_i is the second-order gradient statistic of the loss function, both expressed as

$$g_i = \partial_{\widehat{y_i}^{(t-1)}} l(y_i, \widehat{y_i}^{(t-1)}), h_i = \partial^2_{\widehat{y_i}^{(t-1)}} l(y_i, \widehat{y_i}^{(t-1)}) \tag{14}$$

Because the constant term does not affect the optimisation results, the residuals of the previous iteration are omitted, and the objective function can be modified as follows:

$$\widetilde{Obj}^{(t)} \simeq \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{15}$$

Bringing the defined loss function into the objective function and given that the model is a combination of leaf nodes, the objective function can be expressed as

$$\widetilde{Obj}^{(t)} \simeq \sum_{i=1}^{n} [(\sum_{i \in I_j} g_i)\omega_j + \frac{1}{2}(\sum_{i \in I_j} h_i + \lambda)\omega_j^2] + \gamma T \tag{16}$$

where I_j is the set of leaf nodes. One can simplify the objective function and obtain the optimal solution for ω by defining $G_j = \sum_{i \in I_j} g_i$ and $H_j = \sum_{i \in I_j} h_i$ as follows:

$$\omega_j^* = -\frac{G_j}{H_i + \lambda}, \quad \widetilde{Obj}^{(t)} = -\frac{1}{2}\sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{17}$$

B. FEATURE SELECTION AND DATABASE CREATION

The idea of this study is to build the XGBoost model to solve the multifactorial problem of monitoring driving stress levels in inner-city driving. The model training dataset needs to be built based on data from a real driving task, which includes various potential variables associated with the predicted stress levels, such as driving behaviour, driving environment, and route familiarity. The two methods of quantifying the driving environment, full region and subregion, produce different eigenvalues. After considering the opinions of all the participants, the study determines the initial eigenvalues of the model for each of the two cases, and finally rejects the uncorrelated eigenvalues according to the Pearson correlation coefficient results of the initial eigenvalues and stress levels. Tables 3 and 4 include the lists of the finalised eigenvalues under the segmentation method for the whole region and

TABLE 3. Summary of model characteristics for full regional segmentation.

Class	Category	Features	Description	Number
Classification criteria	Section type	SecType	Type of road section, where 1 is a straight line section and 2 is an intersection section	1
Input attributes	Running state	Acceleration	The resulting acceleration measured by the BIOPACMP160 accelerometer, including X- and Y-axes	2
		Speed	Vehicle running speed values acquired by YOUJIA i-box	1
	Driving environment	Proportion	Proportion of visual scene elements in images extracted from driving videos, including Road, Sidewalk, Building, Wall, Fence, Vegetation, Sky, Person, Car, Bus, and Bicycle	11
		DiffProportion	The difference between the proportion of visual scene elements in the current frame and the average proportion in the previous six frames, including Road, Sidewalk, Building, Wall, Fence, Vegetation, Sky, Person, Car, Bus, and Bicycle	11
		Movetype	Number of types of moveable scene elements	1
	Route familiarity	Familiarity	Number of trials on the same route in the experiment	1
Output target	Driving stress level	ClassifiLevel	Driving stress levels according to HR and SCR, where low stress is 1, medium stress is 2, and high stress is 3	1

TABLE 4. Summary of model characteristics for subregional segmentation.

Class	Category	Features	Description	Number
Classification criteria	Section type	SecType	Type of road section, where 1 is a straight line section and 2 is an intersection section	1
Input attributes	Running state	Acceleration	The resulting acceleration measured by the BIOPACMP160 accelerometer, including X- and Y-axes	2
		Speed	Vehicle running speed values acquired by YOUJIA i-box	1
	Driving environment	Proportion	The proportion of visual scene elements in images extracted from driving videos, including Road (All), Sidewalk (Right), Building (All, Up), Wall (All), Fence (Left, Right), Vegetation (All, Up), Sky (All, Up), Person (Left, Centre, Right, Bottom), Car (Left, Centre, Right, Bottom), Bus (Left, Centre, Right, Bottom), and Bicycle (Left, Centre, Right, Bottom)	27
		DiffProportion	The difference between the proportion of visual scene elements in the current frame and the average proportion in the previous six frames, including Road (All), Building (All, Up), Wall (All), Vegetation (All, Up), Sky (All, Up), Person (Centre, Right, Bottom), Car (Centre, Right, Bottom), Bus (Centre, Right, Bottom), and Bicycle (Centre, Right, Bottom)	20
		Movetype	Number of types of moveable scene elements (Centre, Right, Bottom)	3
	Route familiarity	Familiarity	Number of trials on the same route in the experiment	1
Output target	Driving stress level	ClassifiLevel	Driving stress levels according to HR and SCR, where low stress is 1, medium stress is 2, and high stress is 3	1

subregion, respectively. The number of eigenvalues determined for the full regional segmentation is 27 and that for the subregional segmentation is 55.

After determining the initial eigenvalues of the model, the data from the real driving task were sorted and summarised according to a sampling rate of 6 samples/s, and the resulting database of both types consisted of 239,724 data samples. The degree of change in operating conditions and driving environment varies greatly between straight and intersection sections, and the driver's psychological expectations of changes in operating conditions and driving environment are also very different between the two types of sections. Therefore, separately building different XGBoost models based on road segment type may improve the accuracy of model classification and also investigate the degree of influence of each factor on driving stress on straight and intersection segments. Ultimately, the initial databases were divided into six different categories, with 239,724 samples for All and Sub, 203,376 samples for All_straight and Sub_straight, and 36,348 samples for All_turn and Sub_turn.

C. VALIDATION METHODS

This study ultimately classified driving stress into three classes (high, medium, and low), and the classification task generally used overall accuracy (*Acc*), sensitivity (*Sen*), and precision (*Pre*) to evaluate the performance of the model's multi-classification task. The specific calculation of the three indicators is as follows:

$$Sen = TP/(TP + FN) \times 100\% \quad (18)$$

$$Pre = TP/(TP + FP) \times 100\% \quad (19)$$

$$Acc = (TP + TN)/(TP + FN + TN + FP) \times 100\% \quad (20)$$

where *TP* (true positive) indicates that samples belonging to that stress level are assigned to that class, *FP* (false positive) indicates that samples not belonging to that stress level are assigned to other classes, *TN* (true negative) indicates that samples belonging to that stress level are assigned to other classes, and *FN* (false negative) indicates that samples not belonging to that stress level are assigned to that class. In this study, which is a multi-classification task, the

classification task is split into three two-classification tasks using the macro-average method, and the mean value of each evaluation indicator is calculated as the evaluation indicator of the model, i.e., $MAC - Acc$, $Mac - Sen$, and $Mac - Pre$.

D. MODEL TRAINING

The model is built using Scikit-learn, an open source tool library for machine learning, and according to the theory of XGBoost, the proper setting of hyper-parameters in the model is crucial to the training performance of the XGBoost model, including max_depth, learning_rate, n_estimators, objective, and gamma. Here, max_depth is the maximum depth of the tree, which affects the overfit or underfit of the model to some extent; learning_rate determines the learning rate of the model. n_estimators is the total number of iterations, i.e., the number of decision trees; objective defines the learning task of the model, and as this study is a three-classification task, this parameter is set to multi:softmax, while num_class is set to 3; the subsample refers to the proportion of databases used in training each tree, which can easily lead to overfitting or underfitting of the model; and gamma is the coefficient of the penalty term, which is the decreasing value of the minimum loss function required to split a node, and the larger the value, the more conservative the algorithm is. The study determined the initial values of the hyperparameters of XGBoost based on the amount of data from the training task and the eigenvalues. Table 5 lists the initial values for the different models.

After setting the initial values of the hyperparameters of the XGBoost algorithm, the aggregated data set is divided into a training set and a test set in the ratio of 0.8 and 0.2, and a 10-fold cross-validation is used on the training set to obtain the optimal parameters of the model. In 10-fold cross-validation, the training set is divided into ten equally sized parts, nine of which are used for training the model and one for validation, and the process is repeated ten times. The accuracy of the final classification task is the performance of the model on the test set. The hyper-parameters are then fine-tuned according to the evaluation indicators of the reference model to identify the hyper-parameters that are critical to improving the model performance. It is first studied to determine the fuzzy range of each hyper-parameter based on an equal difference coarse search for the hyper-parameters based on a large range of initial values. This is followed by a detailed search and calculation of all possible combinations of hyperparameters to determine the best estimate of all hyper-parameters. Not all eigenvalues have a positive effect on the performance of the classification task, and after completing a single model optimisation, the XGBoost model will calculate the importance of each eigenvalue to the model, and the eigenvalues with the lowest impact on the model performance will be eliminated in the next iteration. Through self-screening of model eigenvalues, the study can find the implied best performance of the model and filter out all the eigenvalues that affect the model, which helps to study the mechanism of driving stress.

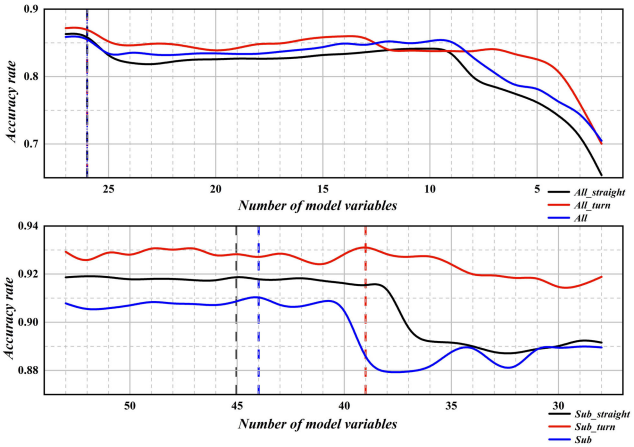


FIGURE 5. Best model accuracy for each iteration of the xgboost model.

V. RESULTS AND DISCUSSION

A. TRAINING RESULTS

The classification performance of the model is not directly proportional to the number of eigenvalues. By iterating the model and filtering the eigenvalues, the study found the best performance for four cases, and the range of hyperparameters used for model training during the iterations is presented in the table. Figure 5 shows the best performance of the three models under different number of eigenvalues. It can be observed that all three models use 26 eigenvalues to achieve the best performance in the case of full regional segmentation, and the precision of All, All_straight, and All_turn reach 86.12%, 86.49%, and 87.39%, respectively. In the subregional segmentation case, Sub, Sub_straight, and Sub_turn used 44, 45, and 39 eigenvalues to achieve the best performance of the model, respectively, with an accuracy of 91.18%, 91.95%, and 93.25% for all three.

B. MODEL PERFORMANCE

Comparing the two segmentation cases, it was found that the model in the case of subregional segmentation based on the distribution of driver attention showed better classification performance compared to the full region segmentation. However, regardless of the segmentation method, if the data can be classified according to the type of road segment in advance, the classification accuracy of the segmented model is higher than that of the original model. At the same time, the classification accuracy of the intersection segment model is higher than that of the linear segment, which may be related to the greater degree of variation in vehicle operation data, driving environment data, and physiological data in the intersection segment. Table 6 presents the $Mac - Acc$(M-A), $Mac - Sen$(M-S), and $Mac - Pre$(M-P) values of the two original models under the full regional segmentation and the subregional segmentation, along with the four models after the segmentation by road segment type. The highest-performing models have M–A, M–S, and M–P values of 93.25%, 89.78%, and 91.43%, respectively, which

TABLE 5. Range of initial and finalised optimal values of hyper-parameters per iteration for different models.

Name	Initial value				Range of optimum values			
	All_straight	All_turn	Sub_straight	Sub_turn	All_straight	All_turn	Sub_straight	Sub_turn
max_depth	10	5	10	5	10–20	5–15	10–30	10–20
learning_rate	0.01	0.01	0.01	0.01	0.01–0.03	0.01–0.02	0.01–0.05	0.01–0.02
n_estimators	100	50	200	100	200–400	50–200	200–600	100–300
subsample	0.6	0.6	0.6	0.6	0.3–0.7	0.5–0.6	0.3–0.7	0.5–0.6
gamma	0.01	0.01	0.01	0.01	0.01–0.02	0.01	0.01–0.02	0.01

TABLE 6. Classification performance of XGBoost and other common machine learning methods.

Type of model	Performance indicators	XGBoost	SVM	RF	GBDT
All	Mac–Acc	86.12	66.60	80.54	83.39
	Mac–Sen	77.09	42.28	67.23	73.47
	Mac–Pre	85.36	58.97	82.84	82.12
All_straight	Mac–Acc	86.49	67.00	79.62	83.47
	Mac–Sen	76.29	42.58	65.32	73.12
	Mac–Pre	85.17	60.70	82.98	82.96
All_turn	Mac–Acc	87.39	67.73	85.55	85.88
	Mac–Sen	82.44	52.49	78.42	78.71
	Mac–Pre	85.49	62.46	83.54	83.13
Sub	Mac–Acc	91.18	71.46	86.87	87.20
	Mac–Sen	84.13	50.89	82.50	82.79
	Mac–Pre	90.25	67.01	86.32	85.91
Sub_straight	Mac–Acc	91.95	71.44	82.32	84.96
	Mac–Sen	85.58	50.90	69.14	72.01
	Mac–Pre	90.54	67.14	86.26	86.76
Sub_turn	Mac–Acc	93.25	76.71	88.40	88.84
	Mac–Sen	89.78	69.15	84.81	85.60
	Mac–Pre	91.43	73.23	87.73	87.26

is a satisfactory result. We also compare the performance of the XGBoost method with three of the more commonly used machine learning methods for stress predictive modelling, namely, SVM, random forest (RF), and gradient boosting decision tree (GBDT). Table 6 also lists the M–A, M–S, and M–P values of the SVM, RF, and GBDT for different data sets for both full regional segmentation and subregional segmentation. The XGBoost model shows the best performance for the classification task of driver's driving stress level.

To further investigate the performance of this model for driving stress monitoring, we compared the performance of this method with that of previous studies. As presented in Table 7, the XGBoost algorithm was used in this study to detect the driver's driving stress, and the performance indicators of accuracy, sensitivity, and precision for the model in the subregional segmentation case were 91.79%–93.73%, 84.42%–90.2%, and 91.01%–91.98%. Regarding the other studies, Healey *et al.* and Singh *et al.* used only physiological data to detect driving stress, classifying the driving stress into 4 and 3 classes, respectively, both with less accuracy than that in our method [40], [58]. Lanatà *et al.* and Rigas *et al.* used physiological data as the primary data source while also considering data of operational conditions or environmental data, and achieved a classification accuracy of 91.33% and 86%, respectively, but both methods were also less accurate than our method [6], [23]. The model performance of the remaining two methods is slightly higher than that of the present study, but the method proposed by Rigas *et al.*, although it can reach 96% accuracy, is only suitable for detecting two stress levels (stress, no stress) [32]. Rastgoo *et al.* used the long short-term memory (LSTM) model to detect the driver's driving stress with high accuracy, taking into account the physiological parameters, vehicle operating conditions, and driving environment [34]. Changes in physiological data are the most direct expression of changes in the driver's stress. Previous methods have mostly used only physiological data or used physiological data as the main data source for models, but because physiological sensors have not been widely used in vehicles, the over-reliance on psychological data to play an active role in the driving stress classification task has prevented the research results from being applied to daily driving tasks. In this study, an attempt was made to monitor the driver's driving stress level using only data from both the vehicle operating conditions and the driving environment in two directions, and the model achieved a better performance without using physiological data.

C. IMPORTANCE OF VARIABLES

Based on the F-weight score of the eigenvalues in the model, we can understand the ranking of importance of each eigenvalue for the model. The F-weight score for each eigenvalue is the total number of times the XGBoost model selects that eigenvalue to generate leaf nodes during training. Figure 6 shows the ranking of importance of the different eigenvalues in the different XGBoost models. It can be observed that the eigenvalues in terms of vehicle operating conditions are of higher importance for all the models, indicating that driving stress is more affected by the vehicle running conditions

TABLE 7. Comparison between present and previous studies.

Authors	Data sources			Performance	Number of classes
	Physiological data	Running condition	Environmental data		
Healey et al.(2000) [58]	ECG, EDA and RSP signals	-	-	Accuracy:88.6%	4 stress class(low, medium, high,very high)
Singh et al.(2013) [40]	PPG, EDA, HRV,RSP signals	-	-	Sensitivity: 88.83% Precision: 89.23%	3 stress class(low, medium, high)
Lanatà et al.(2015) [23]	ECG, EDA, RSP	Vehicle dynamics data	-	Accuracy: 91.33% Sensitivity: 92.33%	3 stress class (no stress, stress level 1,stress level 2)
Rigas et al.(2011) [6]	ECG, EDA, RSP	-	Environmental data	Accuracy: 86%	2 stress class(no stress-stress)
Rigas et al.(2012) [32]	ECG, EDA, RSP	Vehicle dynamic data	-	Accuracy: 96%	2 stress class(no stress-stress)
Rastgoo et al.(2019) [34]	ECG	Vehicle dynamic data	Environmental data	Accuracy: 92.8% Sensitivity: 94.13% Precision: 95.00%	3 stress class(low, medium, high)
Our work	-	Vehicle dynamics data	Environmental data	Accuracy: 91.18%–93.25% Sensitivity: 84.13%–89.37% Precision: 90.25%–91.34%	3 stress class(low, medium, high)

than by changes in the driving environment. In particular, acceleration is more important to the model than running speed, indicating that the degree of variation in running speed is more likely to affect driving stress than the running speed itself. In addition, changes in the transverse acceleration, Accelerate_X, are more likely to affect the driver's stress values than the vehicle's longitudinal acceleration, Accelerate_Y. Route familiarity appeared to have less effect on driving stress, contrary to post-experiment feedback in which participants widely believed that route familiarity had a greater effect on driving stress. We believe that the designed experimental scheme using seven consecutive, continuous trials to familiarise drivers with the route may have accelerated the route familiarisation process and contributed to this result.

In the full regional segmentation case, the proportion of the three immovable scene elements ('sky', 'plant', and 'building') that are important to the model occupy the top three positions in terms of driving environment factors. However, the proportion of vehicle class elements in the Centre and Left regions of the visual scene is more important to the model than in the corresponding model for the full regional segmentation case, and the accuracy of the model is also significantly improved. This indicates that the study of segmentation based on attentional distribution helps the model to refine more important variables.

The rankings of importance of the eigenvalues were not identical for the straight line and intersection segments. The regions within the visual field of the straight line segment were ranked as Up, Right, Left, Centre, and Bottom. Contrary to conventional wisdom, the Centre region, where the driver's attention is most active, is not the most influential region in terms of driving stress. Whereas the intersection segment increases the driver's focus on the Right and Bottom regions while it decreases the focus on the Left region compared to the straight segment. Scene elements such as cars, buildings, sky, and plants are both susceptible to changes in driver stress values for both types of road segments. The difference is that the driver stress status is more likely to be influenced by speed and wall elements on straight segments, whereas drivers on intersection segments are more likely to focus on road and pedestrian elements. We also found another interesting phenomenon in that driver stress conditions seemed to be more influenced by the size of the proportion of common elements, *Proportion*, in the urban landscape in the field of view rather than the difference in their proportions, *DiffProportion*, such as buildings, plants, sky, and vehicles. In contrast, drivers' stress for less common elements in the urban landscape, such as buses and bicycles, seems to be less influenced by their proportional size and is more related to when these elements appear in their vision.

VI. DISCUSSION

Stress has proven to be an important factor in traffic safety, and a growing number of researchers are attempting to detect the driver's stress status [3], [4]. Although previous research results have been able to achieve a high prediction accuracy, the models rely excessively on physiological data and no physiological data acquisition device exists for the vehicle assistance system, making it difficult to apply the previous research to daily driving tasks [1]. Changes in physiological data are the most direct expression of changes in driver's stress condition, but driver's stress is generated by the inability to respond normally to changes in external conditions. In this study, a driver' s driving stress monitoring model is established using only vehicle operation data and driving environment data, and the accuracy of the model without using physiological data is even better than previous studies using physiological data. The performance indicators of this model, accuracy, sensitivity, and precision, reached 91.18%–93.25%, 84.13%–89.37%, and 90.25%–91.34%, respectively. The results of the research can be applied to a new generation of automotive assistance systems to monitor and predict the driver's driving stress using sensors that are commonly found in automobiles, allowing stress monitoring to be widely used in daily driving tasks and improving urban traffic safety. In addition, the study provides

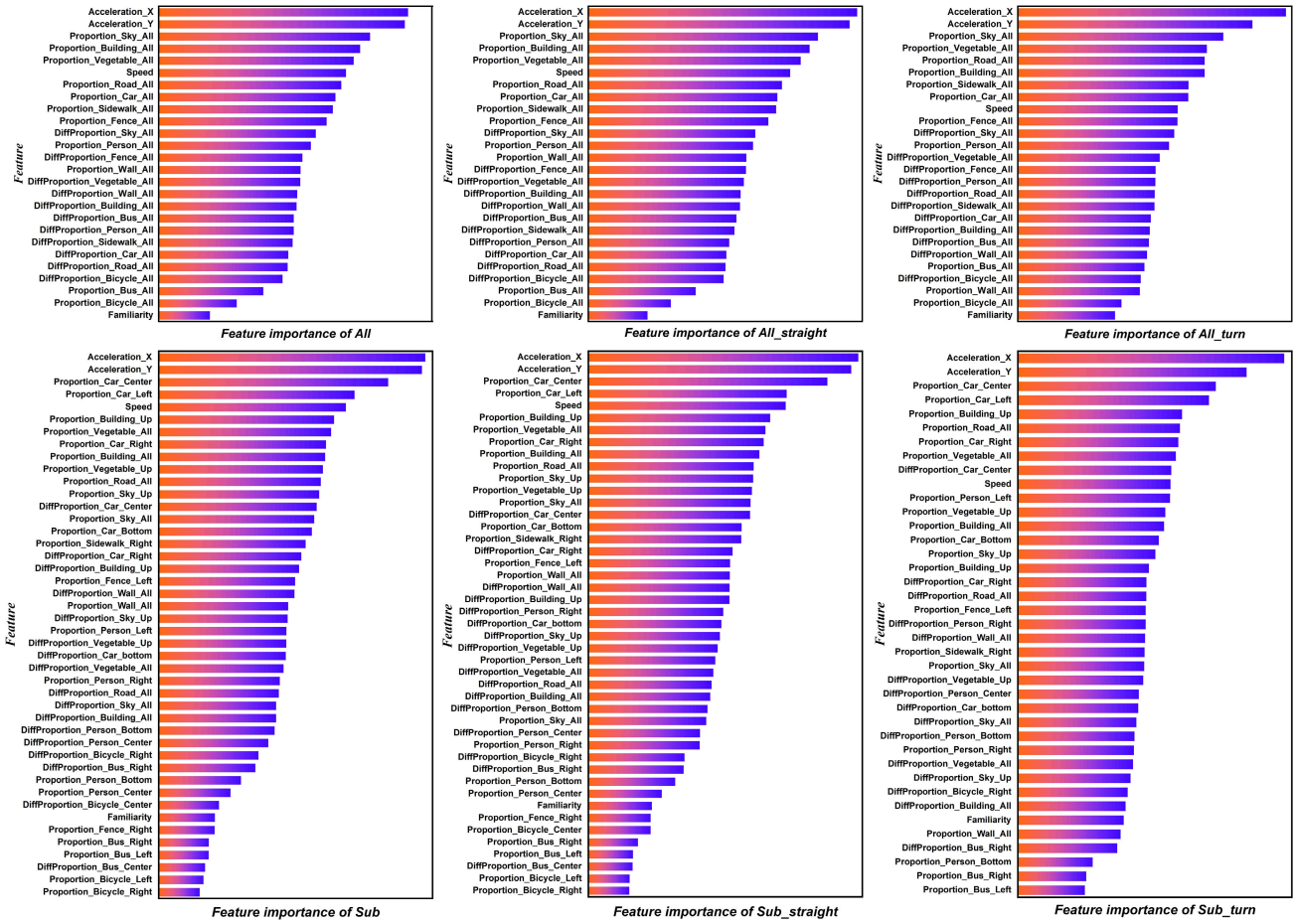


FIGURE 6. Ranking of importance of eigenvalues in different models.

a new perspective on driving stress, exploring the stress status from the perspective of changes in driver perception.

Different urban landscapes can affect drivers' emotions, and a growing number of urban designers are trying to improve people's emotional states through the urban spatial layout and landscape design. This study divides the driving video image into five regions according to the distribution of the driver's attention and quantifies the information of different scene elements as variables of the model for the different regions. Finally, by ranking the importance of the variables, the study not only summarises the ranking of the degree of influence of different regions on driving stress within the field of view, but also summarises the ranking of the influence of different scene elements on driving stress within each region. It is hoped that the results of the research will provide urban designers with recommendations for reducing driving stress and direct their attention to visual regions and visual scene elements that have a higher impact on driving stress and need improvement.

This study still has some shortcomings. On the one hand, the study is based on the image segmentation technique of the DRN model to extract the proportional information of the image scene elements by region to quantify the driving environment, and thus, it is more strict on the requirements of image segmentation accuracy. Lower segmentation accuracy will affect the performance of the driving stress prediction model and the later analysis of the visual scene elements that affect the driver's stress. At the present stage, the semantic segmentation technique can achieve high performance in standard scenarios (with good lighting conditions during daytime), but the semantic segmentation can be affected by unfavourable factors such as nighttime, which may greatly reduce its accuracy. Therefore, this study only monitors daytime driver' s driving stress for the time being. Owing to the driver's limited perception during nighttime driving, external factors that affect driving stress may differ from daytime driving. If future breakthroughs in image semantic segmentation accuracy are achieved at night, we hope to extend this research to nighttime and explore how various scene elements in the nighttime field of view affect the driver's driving stress.

On the other hand, through pre-experiment communication with the participants, we learned that route familiarity had a large impact on driving stress, with the majority of participants reporting that they would experience more stress when

driving on unfamiliar routes. To include route familiarity as a variable in the model, participants were prohibited from being informed of the experimental route prior to the real driving experiment to ensure that they were unfamiliar with the route. The participants were asked to drive seven consecutive trials on the unfamiliar experimental route to familiarise themselves with the route, and the final study used the number of experimental trials as an indicator of route familiarity. Obviously, daily driving does not involve many consecutive trials to familiarise with a route, and there will be some time interval between trials. This means that the driver in this experiment became familiar with the route too rapidly, which resulted in the variable of route familiarity being too low in importance for the model. In future research, we would like to design more rational experiments to investigate the effect of route familiarity on driving stress, taking into account that drivers have a memory curve for the route and may forget details of the route to some extent if the route is not repeated for a certain period of time.

VII. CONCLUSION

The study attempts to move away from the positive effect of psychological data for the improvement in accuracy of previous stress monitoring model studies and proposes a driving stress monitoring model based on driving behaviour, driving environment, and route familiarity. The results show that the performance of the subregional segmentation model based on the distribution of driver attention is better than that of the full region segmentation model. The performance indicators of accuracy, sensitivity, and precision are 91.18%–93.25%, 84.13%–89.37%, and 90.25%–91.34%, and the performance of the classified model is better than that of the original model if the database can be classified by type of road sections beforehand. A comparison with three other mainstream machine learning methods and previous traditional monitoring models shows that the XGBoost model is suitable for driver' s driving stress monitoring and its classification performance exceeds that of most traditional models without using physiological data. We believe that the results of the study could make large-scale application of stress detection in daily driving a possibility. The importance of the eigenvalues for optimal performance is presented at the end of the article, and the study summarises the effects on driving stress of different vision regions and different vision elements within the regions. We hope that the results of our research will provide urban designers with recommendations for reducing driving stress and will direct their attention to those visual regions and visual scene elements that have a higher impact on driving stress and need improvement.

REFERENCES

[1] H. Rahman, M. U. Ahmed, S. Barua, and S. Begum, "Non-contact-based driver's cognitive load classification using physiological and vehicular parameters," *Biomed. Signal Process. Control*, vol. 55, Jan. 2020, Art. no. 101634, doi: 10.1016/j.bspc.2019.101634.

[2] T. G. Brown, M. C. Ouimet, M. Eldeb, J. Tremblay, E. Vingilis, L. Nadeau, J. Pruessner, and A. Bechara, "Personality, executive control, and neurobiological characteristics associated with different forms of risky driving," *PLoS ONE*, vol. 11, no. 2, Feb. 2016, Art. no. e0150227, doi: 10.1371/journal.pone.0150227.

[3] Z. Halim and M. Rehan, "On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning," *Inf. Fusion.*, vol. 53, pp. 66–79, Jan. 2020, doi: 10.1016/j.inffus.2019.06.006.

[4] W. Hadi, N. El-Khalili, M. AlNashashibi, G. Issa, and A. A. AlBanna, "Application of data mining algorithms for improving stress prediction of automobile drivers: A case study in jordan," *Comput. Biol. Med.*, vol. 114, Nov. 2019, Art. no. 103474, doi: 10.1016/j.compbiomed.2019.103474.

[5] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.

[6] G. Rigas, Y. Goletsis, P. Bougia, and D. Fotiadis, "Towards driver's state recognition on real driving conditions," *Int. J. Veh. Technol.*, vol. 2011, Dec. 2011, Art. no. 617210, doi: 10.1155/2011/617210.

[7] R. H. Matsuoka and R. Kaplan, "People needs in the urban landscape: Analysis of landscape and urban planning contributions," *Landscape Urban Planning*, vol. 84, no. 1, pp. 7–19, Jan. 2008, doi: 10.1016/j.landurbplan.2007.09.009.

[8] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet, "Multi-digit number recognition from street view imagery using deep convolutional neural networks," 2013, *arXiv:1312.6082*. [Online]. Available: http://arxiv.org/abs/1312.6082

[9] P. Intini, P. Colonna, and E. Olaussen Ryeng, "Route familiarity in road safety: A literature review and an identification proposal," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 62, pp. 651–671, Apr. 2019, doi: 10.1016/j.trf.2018.12.020.

[10] C. D. Katsis and G. G. Rigas, "Emotion recognition in car industry," *Emotion Recognit.*, vol. 16, pp. 515–544, Feb. 2015, doi: 10.1002/9781118910566.ch20.

[11] I. J. Kopin, G. Eisenhofer, and D. Goldstein, "Sympathoadrenal medullary system and stress," *Mech. Phys. Emotional Stress.*, vol. 4, pp. 11–23, Dec. 1988, doi: 10.1007/978-1-4899-2064-5_2.

[12] E. Gulian, G. Matthews, A. I. Glendon, and D. R. Davies, "Dimensions of driver stress," *Ergonomics*, vol. 32, no. 6, pp. 585–602, 1989, doi: 10.1080/00140138908966134.

[13] L. Dorn and G. Matthews, "Prediction of mood and risk appraisals from trait measures: Two studies of simulated driving," *Eur. J. Personality*, vol. 9, no. 1, pp. 25–42, 1995, doi: 10.1002/per.2410090103.

[14] M. N. Rastgoo, B. Nakisa, A. Rakotonirainy, V. Chandran, and D. Tjondronegoro, "A critical review of proactive detection of driver stress levels based on multimodal measurements," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–35, Jan. 2019, doi: 10.1145/3186585.

[15] T. Lajunen and H. Summala, "Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 6, no. 2, pp. 97–107, 2003, doi: 10.1016/S1369-8478(03)00008-1.

[16] L. Eboli, G. Mazzulla, and G. Pungillo, "The influence of physical and emotional factors on driving style of car drivers: A survey design," *Travel Behav. Soc.*, vol. 7, pp. 43–51, Oct. 2017, doi: 10.1016/j.tbs.2017.02.001.

[17] M. Laila Martinussen, M. Müller, and G. Carlo Prato, "Assessing the relationship between the driver behavior questionnaire and the driver skill inventory: Revealing sub-groups of drivers," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 26, pp. 82–91, Aug. 2014, doi: 10.1016/j.trf.2014.06.008.

[18] G. Matthews, A. Tsuda, G. Xin, and Y. Ozeki, "Individual differences in driver stress vulnerability in a Japanese sample," *Ergonomics*, vol. 42, no. 3, pp. 401–415, Mar. 1999, doi: 10.1080/001401399185559.

[19] T. Lajunen and H. Summala, "Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments," *Personality Individual Differences*, vol. 19, no. 3, pp. 307–318, 1995, doi: 10.1016/0191-8869(95)00068-H.

[20] T. Özkan, T. Lajunen, J. E. Chliaoutakis, D. Parker, and H. Summala, "Cross-cultural differences in driving skills: A comparison of six countries," *Accident Anal. Prevention*, vol. 38, no. 5, pp. 1011–1018, Sep. 2006, doi: 10.1016/j.aap.2006.04.006.

[21] M. Elgendi and C. Menon, "Machine learning ranks ECG as an optimal wearable biosignal for assessing driving stress," *IEEE Access*, vol. 8, pp. 34362–34374, 2020, doi: 10.1109/ACCESS.2020.2974933.

[22] Z. H. Khattak, M. D. Fontaine, and R. A. Boateng, "Evaluating the impact of adaptive signal control technology on driver stress and behavior using real-world experimental data," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 58, pp. 133–144, Oct. 2018, doi: 10.1016/j.trf.2018.06.006.

[23] A. Lanata, G. Valenza, A. Greco, C. Gentili, R. Bartolozzi, F. Bucchi, F. Frendo, and E. P. Scilingo, "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1505–1517, Jun. 2015, doi: 10.1109/TITS.2014.2365681.

[24] C. D. Katsis, N. Katertsidis, G. Ganiatsas, and D. I. Fotiadis, "Toward emotion recognition in car-racing drivers: A biosignal processing approach," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 38, no. 3, pp. 502–512, May 2008, doi: 10.1109/TSMCA.2008.918624.

[25] K. Soman, A. Sathiya, and N. Suganthi, "Classification of stress of automobile drivers using radial basis function kernel support vector machine," in *Proc. Int. Conf. Inf. Commun. Embedded Syst. (ICICES)*, Feb. 2014, pp. 1–5, doi: 10.1109/ICICES.2014.7034000.

[26] T. Yamakoshi, K. Yamakoshi, S. Tanaka, M. Nogawa, S. B. Park, M. Shibata, Y. Sawada, P. Rolfe, and Y. Hirose, "Feasibility study on driver's stress detection from differential skin temperature measurement," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 1076–1079, doi: 10.1109/IEMBS.2008.4649346.

[27] M. Pedrotti, M. A. Mirzaei, A. Tedesco, J.-R. Chardonnet, F. Mérienne, S. Benedetto, and T. Baccino, "Automatic stress classification with pupil diameter analysis," *Int. J. Hum.-Comput. Interact.*, vol. 30, no. 3, pp. 220–236, Mar. 2014, doi: 10.1080/10447318.2013.848320.

[28] M. Paschero, G. Del Vescovo, and L. Benucci, "A real time classifier for emotion and stress recognition in a vehicle driver," in *Proc. IEEE Int. Symp. Ind. Electron.,* Hangzhou, China, 2012, pp. 1690–1695, doi: 10.1109/ISIE.2012.6237345.

[29] R. Fernandez and W. Rosalind Picard, "Prediction of mood and risk appraisals from trait measures: Two studies of simulated driving," *Speech Commun.*, vol. 40, no. 1, pp. 145–159, 2003, doi: 10.1016/S0167-6393(02)00080-8.

[30] H. Boril, P Boyraz, JHL Hansen, "Towards multimodal driver's stress detection," in *Proc. Digit. Signal Process. In-Vehicle Syst. Saf.*, 2012, pp. 3–19, doi: 10.1007/978-1-4419-9607-7_1.

[31] S. Vhaduri, A. Ali, M. Sharmin, K. Hovsepian, and S. Kumar, "Estimating drivers' stress from GPS traces," in *Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, 2014, pp. 1–8, doi: 10.1145/2667317.2667335.

[32] G. Rigas, Y. Goletsis, and D. I. Fotiadis, "Real-time driver's stress event detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 221–234, Mar. 2012, doi: 10.1109/TITS.2011.2168215.

[33] J. D. Hill and L. N. Boyle, "Driver stress as influenced by driving maneuvers and roadway conditions," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 10, no. 3, pp. 177–186, May 2007, doi: 10.1016/j.trf.2006.09.002.

[34] M. N. Rastgoo, B. Nakisa, F. Maire, A. Rakotonirainy, and V. Chandran, "Automatic driver stress level classification using multimodal deep learning," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112793, doi: 10.1016/j.eswa.2019.07.010.

[35] P. Konstantopoulos, P. Chapman, and D. Crundall, "Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving," *Accident Anal. Prevention*, vol. 42, no. 3, pp. 827–834, May 2010, doi: 10.1016/j.aap.2009.09.022.

[36] A. Akbas, "Evaluation of the physiological data indicating the dynamic stress level of drivers," *entific Res. Essays.*, vol. 6, no. 2, pp. 430–439, Feb. 2011.

[37] M. Gutmann, P. Grausberg, and K. Kyamakya, "Detecting human driver's physiological stress and emotions using sophisticated one-person cockpit vehicle simulator," in *Proc. Inf. Technol. Innov. Bus. Conf. (ITIB)*, Oct. 2015, pp. 15–18, doi: 10.1109/ITIB.2015.7355064.

[38] G. Rebolledo-Mendez, A. Reyes, S. Paszkowicz, M. C. Domingo, and L. Skrypchuk, "Developing a body sensor network to detect emotions during driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1850–1854, Aug. 2014, doi: 10.1109/TITS.2014.2335151.

[39] M. Yamaguchi, J. Wakasugi, and J. Sakakima, "Evaluation of driver stress using biomarker in motor-vehicle driving simulator," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 1834–1837, doi: 10.1109/IEMBS.2006.260001.

[40] R. R. Singh, S. Conjeti, and R. Banerjee, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 740–754, Nov. 2013, doi: 10.1016/j.bspc.2013.06.014.

[41] M. Urbano, M. Alam, J. Ferreira, J. Fonseca, and P. Simioes, "Cooperative driver stress sensing integration with eCall system for improved road safety," in *Proc. 17th Int. Conf. Smart Technol.*, Jul. 2017, pp. 883–888, doi: 10.1109/EUROCON.2017.8011238.

[42] J. Wang, J. M. Warnecke, M. Haghi, and T. M. Deserno, "Unobtrusive health monitoring in private spaces: The smart vehicle," *Sensors*, vol. 20, no. 9, p. 2442, Apr. 2020, doi: 10.3390/s20092442.

[43] M. R. Amin and R. T. Faghih, "Robust inference of autonomic nervous system activation using skin conductance measurements: A multichannel sparse system identification approach," *IEEE Access*, vol. 7, pp. 173419–173437, 2019, doi: 10.1109/ACCESS.2019.2956673.

[44] W. Qu, Q. Zhang, W. Zhao, K. Zhang, and Y. Ge, "Validation of the driver stress inventory in China: Relationship with dangerous driving behaviors," *Accident Anal. Prevention*, vol. 87, pp. 50–58, Feb. 2016, doi: 10.1016/j.aap.2015.11.019.

[45] L. J. M. Mulder, "Measurement and analysis methods of heart rate and respiration for use in applied environments," *Biol. Psychol.*, vol. 34, no. 2, pp. 205–236, 1992, doi: 10.1016/0301-0511(92)90016-N.

[46] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008, doi: 10.1109/TPAMI.2008.26.

[47] O. Dehzangi, V. Rajendra, and M. Taherisadr, "Wearable driver distraction identification on-the-road via continuous decomposition of galvanic skin responses," *Sensors*, vol. 18, no. 2, p. 503, Feb. 2018, doi: 10.3390/s18020503.

[48] C. Collet, A. Clarion, M. Morel, A. Chapon, and C. Petit, "Physiological and behavioural changes associated to the management of secondary tasks while driving," *Appl. Ergonom.*, vol. 40, no. 6, pp. 1041–1046, Nov. 2009, doi: 10.1016/j.apergo.2009.01.007.

[49] Y. Han, R. Zhao, B. Wang, N. Xue, C. Wu, and W. Zhang, "The establishment and analysis of the risk event assessment system in urban traffic environment," *IEEE Access*, vol. 6, pp. 51843–51852, 2018, doi: 10.1109/ACCESS.2018.2865859.

[50] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, "Scene parsing through ADE20K dataset," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5122–5130, doi: 10.1109/CVPR.2017.544.

[51] F. Yu, V. Koltun, and T. Funkhouser, "Dilated residual networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 636–644, doi: 10.1109/CVPR.2017.75.

[52] L. Joanne Harbluk, Y. Ian Noy, L. Patricia Trbovich, and M. Eizenman, "An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance," *Accident Anal. Prevention*, vol. 39, no. 2, pp. 372–379, 2007, doi: 10.1016/j.aap.2006.08.013.

[53] T. Horberry, J. Anderson, M. A. Regan, T. J. Triggs, and J. Brown, "Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance," *Accident Anal. Prevention*, vol. 38, no. 1, pp. 185–191, Jan. 2006, doi: 10.1016/j.aap.2005.09.007.

[54] Y. Lu, X. Fu, C. Lu, E. Guo, F. Tang, J. Zhu, and H. Li, "Effects of route familiarity on drivers' psychological conditions: Based on driving behaviour and driving environment," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 75, pp. 37–54, Nov. 2020, doi: 10.1016/j.trf.2020.09.005.

[55] Y. Qu, Z. Lin, H. Li, and X. Zhang, "Feature recognition of urban road traffic accidents based on GA-XGBoost in the context of big data," *IEEE Access*, vol. 7, pp. 170106–170115, 2019, doi: 10.1109/ACCESS.2019.2952655.

[56] X. Gu, Y. Han, and J. Yu, "A novel lane-changing decision model for autonomous vehicles based on deep autoencoder network and XGBoost," *IEEE Access*, vol. 8, pp. 9846–9863, 2020, doi: 10.1109/ACCESS.2020.2964294.

[57] A. B. Parsa, A. Movahedi, H. Taghipour, S. Derrible, and A. Mohammadian, "Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis," *Accident Anal. Prevention*, vol. 136, Mar. 2020, Art. no. 105405, doi: 10.1016/j.aap.2019.105405.

[58] J. Healey and R. Picard, "SmartCar: Detecting driver stress," in *Proc. 15th Int. Conf. Pattern Recognit.*, Sep. 2000, pp. 218–221, doi: 10.1109/ICPR.2000.902898.

YUE LU was born in Shijiazhuang, China, in 1993. He received the B.S. degree in civil engineering from the Changsha University of Science and Technology, Changsha, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, China. His research interests include traffic safety analysis, and intelligent transportation system (ITS).

XINSHA FU gained state council special allowance in 1993, and was exceptionally promoted as a Professor in 1998. He works on highway planning and design, computer aided engineering and design of highway, transportation infrastructure management system, intelligent transportation system, 3s technology, as well as teaching and research of traffic information. He has published over 60 articles and five monographs. He was awarded two second prizes, and seven third prizes of provincial and ministry level awards.

ENQIANG GUO is currently pursuing the Ph.D. degree with the School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, China. His research interests include intelligent transportation system, semantic segmentation, deep metric learning, and attention mechanism in deep learning model.

FENG TANG was born in Huaihua, China, in 1992. He received the B.S. degree in civil engineering from the Changsha University of Science and Technology, Changsha, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Civil and Transportation Engineering, South China University of Technology, Guangzhou, China. His research interests include traffic safety analysis, and intelligent transportation system (ITS).